HHS Public Access

Author manuscript

Int J Neural Syst. Author manuscript; available in PMC 2021 May 01.

Published in final edited form as:

Int J Neural Syst. 2020 November ; 30(11): 2050030. doi:10.1142/S0129065720500306.

Automated Detection of Interictal Epileptiform Discharges from Scalp Electroencephalograms by Convolutional Neural Networks

John Thomas,
School of Electrical and Electronic Engineering Nanyang Technological University Singapore 639798, Singapore

Jing Jin,
School of Electrical and Electronic Engineering Nanyang Technological University Singapore 639798, Singapore

Prasanth Thangavel,
School of Electrical and Electronic Engineering Nanyang Technological University Singapore 639798, Singapore

Elham Bagheri,
School of Electrical and Electronic Engineering Nanyang Technological University Singapore 639798, Singapore

Rajamanickam Yuvaraj,
School of Electrical and Electronic Engineering Nanyang Technological University Singapore 639798, Singapore

Justin Dauwels*,
School of Electrical and Electronic Engineering Nanyang Technological University Singapore 639798, Singapore

Rahul Rathakrishnan,
Division of Neurology, National University Hospital Singapore 119074, Singapore

Jonathan J. Halford,
Department of Neurology, Medical University of South Carolina Charleston, SC 29425, USA

Sydney S. Cash,
Department of Neurology, Massachusetts General Hospital Boston, MA 02114, USA

Harvard Medical School, Boston, MA 02115, USA

Brandon Westover
Department of Neurology, Massachusetts General Hospital Boston, MA 02114, USA

Harvard Medical School, Boston, MA 02115, USA

Abstract

*Corresponding author : jdauwels@ntu.edu.sg.

Conflicts of Interest
The authors have no disclosures to report.

Visual evaluation of electroencephalogram (EEG) for Interictal Epileptiform Discharges (IEDs) as distinctive biomarkers of epilepsy has various limitations, including time-consuming reviews, steep learning curves, interobserver variability, and the need for specialized experts. The development of an automated IED detector is necessary to provide a faster and reliable diagnosis of epilepsy. In this paper, we propose an automated IED detector based on Convolutional Neural Networks (CNNs). We have evaluated the proposed IED detector on a sizable database of 554 scalp EEG recordings (84 epileptic patients and 461 nonepileptic subjects) recorded at Massachusetts General Hospital (MGH), Boston. The proposed CNN IED detector has achieved superior performance in comparison with conventional methods with a mean cross-validation area under the precision–recall curve (AUPRC) of 0.838 ±0.040 and false detection rate of 0.2±0.11 per minute for a sensitivity of 80%. We demonstrated the proposed system to be noninferior to 30 neurologists on a dataset from the Medical University of South Carolina (MUSC). Further, we clinically validated the system at National University Hospital (NUH), Singapore, with an agreement accuracy of 81.41% with a clinical expert. Moreover, the proposed system can be applied to EEG recordings with any arbitrary number of channels.

Keywords

Epilepsy; electroencephalogram; interictal epileptiform discharges; convolutional neural networks; spike detection; deep learning; clinical validation; multi-center study

1. Introduction

Epilepsy is the fourth most common neurological disorder in the world, with about 70 million affected.[1] It is characterized by enduring predisposition to recurrent unprovoked seizures. The routine scalp electroencephalogram (EEG) is widely explored as a fundamental medical test for diagnosis of epilepsy.[2–7] Diagnosis based on ictal or seizure events is time-consuming due to their low frequency of occurrence. Paroxysmal abnormalities in EEG or Interictal Epileptiform Discharges (IEDs) are unique biomarkers of epilepsy.[8] According to the International Federation of Societies for Electroencephalography and Clinical Neurophysiology (IFSECN), IEDs are defined as transients distinguishable from background activity, with a characteristic spiky morphology, typically, but neither exclusively nor invariably, found in interictal EEGs of individuals with epilepsy.[9] These transients play a crucial role in the diagnosis and treatment of epilepsy.[10]

The current gold standard of identifying IEDs is the visual inspection performed by experts. This is a time-consuming, tedious, and subjective process.[6] Therefore, the development of an automated, efficient IED detection is necessary to improve patient care. Various approaches to IED detection have been proposed in the literature, namely, mimetic analysis, parametric modeling, template matching, neural networks, and wavelet-based ones.[3] In Table A.1, we have tabulated the IED detection literature chronologically. It is infeasible to compare the different methods in the literature due to the lack of a standard IED dataset, diverse IED feature definitions, lack of clinical validation, and distinctive performance evaluation characteristics. Most studies in the literature are only validated on a small database of IEDs, e.g. 1491 (19 patients),[11] 469 (eight patients),[12] 101 (one patient),[13] 482

(eight patients),[14,15] etc. Moreover, the computational complexity of the methods is not reported. The development of a robust and reliable automated or semi-automated IED detection remains a necessity, yet challenging problem.

In the recent years, Convolutional Neural Networks (CNNs)[16] have shown to achieve higher efficiency for epilepsy diagnosis and monitoring: seizure detection,[17,18] EEG classification, [4,5,19] and IED detection.[20–22] Deep networks can detect latent structures or features from the raw data, thereby reducing the dependency on handcrafted features. There is only a limited number of studies investigating the deep learning technique of CNN for IED detection. Le *et al.*[11] have proposed a Deep Belief Network (DBN) for detecting IEDs. Tjepkema-Cloostermans *et al.* developed a CNN-based IED detector that achieved a sensitivity of 47.4% for specificities of 98% (epileptic EEGs) and 99% (nonepileptic EEGs) on a dataset of 100 subjects (with 1815 annotated IEDs).[21] Both have been evaluated only on a small dataset, thereby making it unreliable. Clarke *et al.* employed a two-dimensional (2D) CNN for identifying epileptic events in patients with idiopathic generalized epilepsy (24-h ambulatory outpatient EEG recordings).[22] The system has attained a false positive (FP) rate of 1.16 per minute for a sensitivity of 96.7%.

In this study, we develop a one-dimensional (1D) CNN-based IED detector and compare the performance with the various traditional classification algorithms, namely, *k*-Nearest Neighbors (*k*-NN),[23] Classification and Regression Tree (CART),[24] Support Vector Machine (SVM),[25] Linear Discriminant Analysis (LDA),[26] and Ensemble AdaBoost[27] method (a popular boosting algorithm based on LDA, *k*-NN, and CART). We compare the various classification algorithms by performing five-fold cross-validation on the dataset of 554 subjects (18,164 annotated IEDs) recorded at Massachusetts General Hospital (MGH), Boston. The proposed IED detector has shown superior performance compared to the traditional approaches with a mean false positive rate of 0.2 per minute for a sensitivity of 80%. Further, this study considers EEG data from multiple institutions and also expert annotations from different institutions. We evaluate the proposed system on two other datasets from the Medical University of South Carolina (MUSC), USA, and National University Hospital (NUH), Singapore. For the MUSC and NUH datasets, the IED annotations have been performed independently by different sets of academic and clinical experts. We also evaluate the detector on the EEGs of individual patients and verify whether it detects IEDs in the correct channels. The IED channel localization is essential to aid in determining the epilepsy syndrome as well as help to determine the epileptic foci.[10] As far as we know, such a study has not yet been reported in the literature.

In a preliminary study from 2016, we proposed a 1D CNN for IED detection, achieving an area under the receiver operating characteristic curve (AUC) of 0.947 on a small dataset of five EEGs.[20] In 2018 we reported results for an improved 1D CNN algorithm for IED detection, achieving a false positive rate of 2.38 per minute for a sensitivity of 80%.[4] In that study, we also proposed an automated system to classify IED EEGs versus non-IED EEGs that achieved a five-fold cross-validation accuracy of 83.86% and an AUC of 0.87. In a parallel work to this study, a 2D CNN for IED detection was proposed (SpikeNet), and it was applied for EEG classification that achieved an AUC of 0.98.[5] In this study, we improve the 1D CNN IED detection system and validate it on multiple sizable datasets from different

institutions. This study might be the first cross-institutional assessment of automated IED detection. We assess the sensitivity and false positive rate of the proposed IED detector, and also verify whether it is able to reliably localize IEDs on the scalp.

The rest of the paper is organized as follows: we describe the EEG datasets in Sec. 2.1, the data preparation and preprocessing in Sec. 2.2, and the different methodologies adopted to train and evaluate the IED detector in Secs. 2.3 and 2.4. In Sec. 3, we illustrate the results and observations. Further, in Sec. 4, we discuss the effectiveness of the CNN detector in comparison with the literature. In Sec. 5, we summarize the study and propose the future scope of the study.

2. Methods

2.1. Clinical scalp EEG

We analyzed 554 scalp EEG recordings from 545 subjects (84 epileptic patients with annotated IEDs and 461 nonepileptic EEGs). The data were recorded during clinical care (see Table 1) according to the International 10–20 electrode system at MGH, Boston, USA. The study was approved by the MGH Institutional Review Board (IRB) with a waiver of informed consent. The epileptic dataset contains data from 43 males (aged 35.2±27.2 years) and 41 females (aged 37.1±28.21 years). A total of 14,170 IED events (along with the channel labels of occurrence) are annotated by two neurologists, independently. The annotation was performed with the aid of Neuro-Browser (NB) software.[28] We consider the waveforms with majority agreement (both the annotators) as IEDs. The IEDs were annotated in an exhaustive manner, i.e. the annotators have tried to annotate all the IEDs in the EEGs by reviewing the complete EEG recording. Consequently, a single IED event could occur at multiple channels. We have 14,170 IEDs events with a total of 18,164 IEDs. The inter-IED annotation time window was set as 250ms (32 samples), i.e. only IEDs with a separation of 32 samples are annotated.

In addition, we evaluate the performance of the proposed IED detector on two other datasets from the MUSC[29,30] and NUH, Singapore. The MUSC dataset consists of 200 30-s scalp EEG segments extracted from a library of 1000 routine EEGs performed in the MUSC Neurophysiology Laboratory for clinical purposes between 2010 and 2011. These 200 segments consisted of IEDs, difficult-to-interpret IEDs, benign paroxysmal activity, and random EEG segments from nonepileptic EEGs.[29,30] The data were recorded at different sampling frequencies: 200, 250 and 512Hz. This study was performed according to the IRB approval from MUSC and Clemson University. The MUSC dataset was annotated in two phases with the EEGnet software.[31] In Phase 1, 35 annotators viewed EEG segments and indicated the presence of all IEDs using a Web-based EEG annotation system.[29] A total of 573 IED instants were annotated by at least one expert. In Phase 2, 235 selected IEDs (which were marked by at least two experts) were reannotated by 18 experts on a five-point Likert scale as either (1) definitely not an epileptiform discharge, (2) not likely an epileptiform discharge, (3) not sure, could go either way, (4) likely an epileptiform discharge, or (5) definitely an epileptiform discharge.[30]

The NUH data consists of 75 routine scalp EEGs recorded during the routine clinical care at NUH, Singapore (see Table 2). The study was approved by the National Health Group (NHG) Domain-Specific Review Board (DSRB). The data consists of 75 epileptic EEGs with 43 males (aged 58.75±18.38 years) and 32 females (aged 63.85±17.76 years) recorded at a sampling frequency of 256Hz. We have performed the clinical validation of the proposed system with the NUH dataset. The NUH data were evaluated with the IED detector, and the results are cross-evaluated by a clinical neurologist. The detected IED segments are displayed to the clinical neurologist as a 11-s 19-channel EEG segment (typical clinical setting) in a bipolar longitudinal montage with the identified IED at the center. The clinician annotated the segments on a scale of 1–10, with 1–3 indicating a definite misdetection, 4–6 indicating uncertain segment, and 7–10 indicating a highly certain IED.

2.2. Data preparation

We apply the following EEG preprocessing steps: a Butterworth notch filter (fourth-order) of 60Hz (USA)/50Hz (Singapore) to remove the electrical interference, a Butterworth 1-Hz high-pass filter (fourth-order) to remove the DC shift and baseline fluctuations, downsampling to 128Hz, and the Common Average Referential (CAR) montage. We also performed an artifact rejection based on noise statistics (adapted from recording stability analysis[32]). First, we divided the data from each channel into consecutive segments of 1s and computed the root-mean-square (RMS) value for each segment. For each channel data, the segments that exhibited a higher RMS value than a threshold were identified as noisy ones. The noisy segment was suppressed, i.e. the values were made to zero. If more than five-channels of EEG was found to be noisy, then the entire 1s of 19-channel EEG data were discarded. This mechanism helps us to remove localized artifacts such as electrode popping, faulty electrode contacts, undesired glitches, etc. The threshold was set as median +4×standard deviation of the RMS values.

We extracted the IED segments and non-IED segments to train the IED detectors. Each annotated IED was extracted as a 500-ms (64 samples) waveform. The IEDs are annotated at random locations by the neurologists. To ensure the 500-ms window covers the IED and the associated events, we realign the EEG segments containing IED waveforms manually. The background EEG waveforms (or the non-IEDs) were also extracted from the IED-free EEGs as 500-ms waveforms with an overlap of 75%. We considered only nonepileptic EEGs to extract background waveforms since there is a possibility of overlooked epileptiform events. There is a high probability for the IED event to occur at multiple channels. While evaluating the performance of the system, considering the multiple channel occurrences of a single IED event as different data points might skew the performance indices. Consequently, we applied all the annotated IED waveforms for training, while the testing performance indices are computed based on IED time instants.

The IED detectors were evaluated according to the five-fold cross-validation. We divided the entire dataset EEG-wise into five folds at random by keeping the IED count, age, and gender distribution consistent over the different folds. We have also made sure that the multiple recordings from the same patient are added into the same fold. The details regarding the

different folds are presented in Table 3. We have observed the ratio of IED events to non-IED events to be approximately 1:650 based on the MGH dataset annotations. Each fold contains about 90 nonepileptic EEGs; consequently, each testing set contains approximately 1.25 million background events extracted with an overlap of 75%, typically representing the real-world scenario. The entire block diagram of the proposed approach is illustrated in Fig. 1 (training) and Fig. 2 (evaluation).

2.3. CNN-based IED detector

CNNs are multi-layered feed-forward neural networks.[16] The weights of the network are updated through the process of error backpropagation. It consists of a combination of the following types of layers: the input layer, convolutional layer, rectified linear unit (ReLU) layer, pooling layer, and fully connected layer. Multiple parameters can be optimized for the CNN network, which leads to many possible configurations. The major parameters for a CNN network design include the number of convolutional layers, number of convolution filters, the dimension of convolution filters, number of pooling layers, number of fully connected layers, activation function, number of hidden layer neurons, dropout probability, and number of training iterations. The CNN was implemented with TensorFlow 1.2.1[33] utilizing a K80 Tesla graphical processing unit (GPU) on Ubuntu 16.04.

The CNN IED detector is trained and evaluated at the waveform level, i.e. each IED detector is trained to identify whether a single-channel EEG segment (1×64 segments) is an IED. The set of extracted IED and background waveforms are applied to train the CNN IED detector. Initially, the convolution operation is performed by convolving the EEG segment with a 1D filter to capture the temporal information of the EEG signal and passed through the output function to form the feature map. Here, the output function is selected as the ReLU. Next, a max-pooling layer is applied for dimensionality reduction. The generated features from the convolutional layers are applied to the fully connected hidden layers. Finally, the CNN output layer was mapped to [0,1] by applying a softmax layer, with “1” indicating an IED detection with high confidence (certainty). An illustration of the CNN model applied in this study is provided in Fig. 3.

In order to prevent negative class overfitting due to the class imbalance, we applied balanced training, i.e. the IED-to-background ratio was maintained as 1:1. We also employ dropout and batch processing to optimize the CNN network and to avoid overfitting. In dropout, the weights of a certain percentage of the fully connected layer were dropped at each training iteration to prevent potential overfitting. The two classes, IED and non-IED waveforms, are heavily skewed. In batch processing, the training class ratio is maintained as 1:1, but the training set is continuously replaced after a prefixed number of iterations (10 iterations). We also applied a negative sampling technique to improve the IED training process, i.e. we extracted the background waveforms that were most similar to the IEDs. The entire background set contains approximately 99% of highly certain or easily separable waveforms. We employed a CNN IED detector trained on random backgrounds to separate the IED-like negative samples or “false positive” backgrounds. Later the training was performed based on the extracted FP background set.

The hyperparameters of CNN and the training termination criteria were optimized in a brute force manner with nested cross-validation on the training dataset: 75% for training and 25% for validation. The CNN training termination criteria are dependent on the validation error. As a result, we need a separate validation set in order to evaluate the different sets of hyperparameters. Consequently, we divide the validation set into two: one for training termination criteria (Validation Set 1), and the other for the hyperparameter selection (Validation Set 2). For each fold, the network was trained until the network attained a particular patience level. The patience-level threshold was set at 25. Each time when there was no improvement in the validation loss over five consecutive iterations (the improvement threshold was set as the learning rate 10E−4), the patience variable was incremented. Later, we evaluate the performance of the trained CNN on Validation Set 2. This process is repeated for the different sets of CNN hyperparameters, and the set that produces the best performance is chosen as the optimized hyperparameter set. Table 4 details the different optimized parameters applied for CNN. We consider the cross-entropy as the error function for the training and validation of the CNN architecture.

After applying the CNN IED detector, we obtain 19 outputs (different EEG channels) for a single EEG time instant. Next, we combine these 19 outputs by considering the maximum of the 19-channel outputs to produce a single output value. When the output value is above a threshold, an IED occurrence is identified. In our analysis, the training is performed on all IEDs in the training set (including the multiple occurrence of an IED), whereas the evaluation is only performed on the IED events.

2.4. Performance assessment

We compare the CNN IED detector with traditional classifier-based IED detectors: *k*-NN,[23] CART,[24] SVM,[25] LDA,[26] and Ensemble AdaBoost[27] method (LDA, *k*-NN, and CART). We apply the same pipeline as well as data cross-validation for evaluating the different classifiers. The different parameters of the traditional classifiers (see Table B.1) are optimized by applying Bayesian optimization[34] (five-fold cross-validation). We investigate the performance of IED detectors based on the measures listed in Appendix C. Since the different classes are heavily imbalanced, we assess the performance of the classifiers by means of area-related measures, namely, AUC and area under the precision–recall curve (AUPRC). At times, area-related measures also could be misleading. Consequently, we evaluate the false positives per minute values for fixed sensitivity values. After evaluating the IED detection performance, we also evaluate whether the IEDs are detected at the correct channels. Specifically, we compare the channel labels annotated by the experts with the channel labels predicted by the IED detector. This was also performed as part of the five-fold cross-validation, where the CNN threshold was chosen to be 0.5. Moreover, we investigate the performance of the IED detector for each EEG; concretely, we report how many of the IEDs were detected by the CNN algorithm.

We also evaluate IED detectors trained on the MGH dataset on the MUSC and NUH datasets. Here, we evaluate the datasets by applying a sliding window of 75%, and no realignment for the IEDs is performed. In the MUSC dataset analysis, we set a threshold of 4 (as annotations 4 and 5 indicate a confirmed IED) on the mean annotator consensus value

on Phase-2 annotations and evaluate the performance of the IED detectors in comparison with the performance of the 35 experts. The detection window was set as 1s, i.e. the detections separated by 1s or less were considered the same detection. In the NUH data analysis, we evaluate the EEGs with the CNN IED detector and randomly extracted IED events with higher certainty (CNN prediction value greater than 0.4). These IEDs are later scored by an expert (Rahul Rathakrishnan) at NUH. Finally, we investigate the relevance of features that defines an IED. We investigate the common morphological features of IEDs (see Fig. 4), i.e. amplitude (peak, rise, and fall amplitudes), duration (rise and fall durations), and slope (rise and fall slopes) features.[35] CNNs have been proved to learn sophisticated features without supervision. Here we apply the time-series EEG waveforms as the input to the CNN. We compare the feature relevance by evaluating the correlation with the CNN output. The relevance was computed for each cross-validation fold separately, and the results were averaged.

3. Results

3.1. CNN IED detector

The five-fold cross-validation AUC and AUPRC results for the IED detectors are presented in Table 5. The AUC values are high for all the IED detectors on the MGH EEG dataset. This is mainly attributed to class imbalance. In such scenarios, AUPRC is more meaningful compared to AUC values. We discard the detectors that have achieved inferior AUPRC values (value less than 0.5) from further analysis.

The IED detectors based on CNN, Ensemble CART, and SVM with Gaussian kernel provided the best AUPRC values. Now we compare the false positives per minute for fixed sensitivity values for further evaluating the three IED detectors (see Table 6). The traditional IED detectors achieved the values of mean false positives per minute of 4.72±5.43 (Ensemble CART) and 2.55±3.48 (SVM), whereas the CNN IED detector exhibited superior performance with a false positive rate of 0.2±0.11 (sensitivity of 80%). We have compared the different methods in the literature with the proposed CNN IED detector and the traditional IED detectors in Fig. 5. The proposed IED detector has exhibited superior performance than most of the methods reported in the literature.

3.2. Patient-wise assessment of the IED detector

We investigate here the sensitivity of the CNN IED detector for each individual EEG (patient) separately. Specifically, in Fig. 6, we show the true positives versus the total number of IEDs (ground truth) for the 93 annotated EEGs. For the great majority of the EEGs, most IEDs were detected by the CNN algorithm (see sensitivities for different EEGs in Fig. 7). For three of the EEGs (with only one annotated IED), the IEDs were not identified. The median error percentage is 1.59%, and the mean error percentage is 12.9% [95% confidence interval (CI), 8.3–17.5%].

3.3. IED channel localization

It is important to reliably determine the location of the IEDs on the scalp. To this end, the IED detector needs to detect the IEDs at the correct locations. So far in our analysis, we

ignored this issue, since the detector aggregates CNN outputs from all channels, and decides based on the maximum of those outputs. Here we investigate whether the CNN outputs at the individual channels match the locations of the IEDs. Specifically, we evaluate whether the CNN IED detector has detected the IEDs at the same channel locations as annotated by the clinical experts. We apply the same five-fold cross-validation analysis (see Sec. 2.2), and we set the CNN IED detector threshold as 0.5 to make a binary decision (IED or non-IED). As main result of this analysis, we report that the proposed IED detector identifies the channel index at which the IED occurs with a sensitivity of 90.67%, i.e. 16,469 IEDs (out of 18,164) are identified at the same channel locations as annotated by the experts. We also investigate the performance of channel identification for each individual EEG (patient) separately. In terms of patientwise assessment, the system achieves a sensitivity of 82.03% (95% CI, 77.02–87.05%).

3.4. Feature evaluation

We evaluated the significance of the morphological characteristics of IEDs by computing their correlation with the CNN output values. Concretely, we compute the Pearson correlation coefficient. The correlation values for the different features are illustrated in Fig. 8. The correlation values for amplitude and slope features are positive. This suggests that the output of the CNN IED detector will be higher for waveforms with high amplitudes, and steep rise and fall slopes. The correlation values for the duration features were negative, i.e. detected IEDs are generally of shorter duration. We observe that the rising slope is the most consistent feature of an IED, as it has the most substantial average correlation and lowest standard deviation values.

3.5. Evaluation on MUSC and NUH datasets

We evaluated the three IED detectors (CNN, SVMG, and Ensemble CART) on the MUSC dataset. The results for the evaluation with mean annotator consensus of 4 are presented in Fig. 9 along with the individual annotator performance. The CNN IED detector achieved a false detection rate of 1.43 per minute for a sensitivity of 80%. The proposed IED detector was observed to be noninferior to most of the annotators. In the evaluation on the NUH dataset, we randomly selected 136 IED events with high certainty (CNN threshold greater than or equal to 0.4). These were cross-evaluated by an expert at NUH (Rahul Rathakrishnan) on a scale of 1–10. In Fig. 10, we present the CNN outputs together with the annotator scores. The proposed system has achieved the following agreement values (for different thresholds) with the expert annotations for the IED events:

- 84.55% (CNN prediction ≥0.4, annotator score ≥4);
- 81.41% (CNN prediction ≥0.5, annotator score ≥5); and
- 77.17% (CNN prediction ≥0.6, annotator score ≥6).

4. Discussion

We have developed an automated IED detector based on CNN, and compared it to standard classifiers including SVMG and Ensemble DT. The different IED detectors were evaluated by performing five-fold cross-validation with the MGH dataset of 554 subjects. Among the

traditional IED detectors, SVM with Gaussian kernel and AdaBoost CART performed better than other classifiers with mean false positive rates of 4.72 per minute and 2.55 per minute (sensitivity of 80%), respectively. But the proposed CNN IED detector has achieved considerably superior false positive rates of 0.2 per minute and 0.92 per minute for the sensitivity thresholds of 80% and 90%, respectively. We evaluated the IED detectors on the MUSC database, and the comparison results were corroborated with the five-fold cross-validation results. Moreover, the CNN IED detector was demonstrated to be noninferior to the neurologists (both academic and clinical) on the MUSC dataset. The IED detector was clinically validated at NUH, Singapore, with an agreement of 81.41% with the consensus of a single clinical expert. Therefore, the proposed system performance was comparable with an uncorrelated expert opinion.

The proposed IED detector has achieved superior performance in comparison with the different methods proposed in the literature (see Fig. 5). Le *et al.* have proposed a DBN for detecting IEDs. Their study is performed on data from 19 subjects with 1491 annotated events. The study employs a feature extraction method in combination with a DBN and has achieved a mean leave-one-subject-out cross-validation (LOSO-CV) sensitivity of 76% for the specificity of 97.24%. The results are low for clinical application even though the testing is performed on an IED-to-background ratio of 1:2.[11] Tjepkema-Cloostermans *et al.* had developed a CNN-based IED detector, which is shown to have achieved a sensitivity of 47.4% for specificities of 98% (epileptic EEGs) and 99% (nonepileptic EEGs) on a dataset of 100 subjects (with 1815 annotated IEDs).[21] The evaluation was reported for a single run, i.e. cross-validation was not performed. Moreover, the study does not report any benchmark results. Scheuer *et al.* have evaluated the Persyst P13 detector on a dataset of 5474 IEDs found within 40 long-term EEGs, 35 of which contained IEDs and achieved a sensitivity of 43.9% for an FP rate of 1.65 per minute.[36] The study also benchmarks three human annotators with a mean pair-wise sensitivity of 51.5% for a false detection rate of 1.99 per minute. Ganglberger *et al.* have demonstrated a similar study comparing the EpiSpike 1.5 algorithm[37] applied in the Enceveis software with ANN-based system on a dataset of 12 long-term EEGs with 5582 annotated IEDs. The ANN-based system performed superior to the EpiSpike algorithm for the same false positive rate of 0.33 per minute (ANN-based system: sensitivity of 44.1% for a precision of 56.2%, EpiSpike: sensitivity of 34.6% for a precision of 58.5%). Jing *et al.* have achieved a 10-fold cross-validation AUC of 0.98 for classifying IED according to the consensus of eight annotators by applying a 2D CNN-based IED detector (SpikeNet[5]). The other performance indices are not reported for the above study, and therefore, a direct comparison cannot be performed. The CNN IED detector proposed in this paper achieved a superior mean five-fold cross-validation AUC of 0.989.

Three methods[13,22,38] have been found to have reported superior performance than our proposed IED detector [see Fig. 5(a)]. Zacharaki *et al.* have evaluated the system only on a small dataset of 101 IEDs from a single EEG,[13] whereas our proposed system has been evaluated on a sizeable dataset of 554 subjects with 18,164 annotated IEDs. Thomas *et al.* have performed the evaluation only on a subset of the EEG segments[38] with IED-to-background ratio of 1:5. In our study, we have considered the complete 30-min EEG recordings with the typical IED-to-background ratio of 1:500. Clarke *et al.* have developed a system for identifying epileptic events in patients with idiopathic generalized epilepsy who

underwent 24-h ambulatory outpatient EEG recording.[22] Our study mainly focuses on developing a generalized IED detector on routine scalp EEG recordings. Our proposed IED detector has achieved a sensitivity of 80% for an FP rate of 0.2 per minute on a sizable database of 14,170 IED events. Also, we have employed one of the largest databases of interictal events for evaluation from multiple centers.

Moreover, a 1D CNN, as proposed in this paper, offers significant advantages over a 2D CNN. First of all, it contains on the order of 10–100 times fewer parameters, hence it requires less data and time to train. Second, since the detector is applied to each channel separately, it is able detect at what channels the IEDs occur, yielding important spatial information about the IED. Such information is harder to obtain from a 2D CNN, since only a single output is provided for a segment of multi-channel EEG, instead of outputs for each individual channel as in 1D CNN. Moreover, the proposed 1D CNN can directly be applied to an arbitrary number of channels, whereas a separate 2D CNN model needs to be trained for every number of channels; this can potentially be achieved by transfer learning. In other words, one would need to train a 2D CNN model for 16, 32, 64, etc. channels. In practical EEG recordings, the signal at a few channels might be distorted, due to a bad contact or other reasons. In the 1D CNN approach, such channels can readily be left out in the detection process, while that is far less obvious or straightforward for 2D CNN. As a last advantage of the 1D CNN approach, we would like to point out that it is trivial to parallelize the 1D CNN IED detector code on GPU, since the 1D CNN at each channel can be implemented on a separate core. Although the 2D CNN can also be implemented in a parallel fashion, it is less trivial.

The proposed CNN IED detector takes approximately 0.744 (95% CI, 0.740–0.748) for preprocessing and 8.180s (95% CI, 8.162–8.198) for CNN evaluation (determined from 100 runs) for evaluating a single 30-min 19-channel routine EEG recording sampled at 128Hz. The above evaluation does not include the data loading overhead. An expert takes approximately 10min to review a 30-min routine EEG with approximately 19 electrodes. By contrast, the CNN IED detector can analyze the EEG within seconds, providing the neurologist valuable information about potential IEDs in the EEG. As a result, the proposed IED detector may speed up the diagnostic process. Moreover, the CNN detector can be operated in real time, which might prove to be useful for analyzing EEG in real time in neuro-ICU settings. The evaluation was performed with an Intel (R) Xeon(R) CPU E5–2630 v4 @ 2.2-GHz CPU using Python v3.5.

This study has a few limitations. The simple artifact rejection technique applied in this study is only capable of removing high-amplitude noise. More sophisticated artifact rejection has to be employed for improving the IED detector. Moreover, we have applied a simple “max rule” for combining the output from the different channels into a single time-instant output. This typically leads to higher sensitivity at the cost of more false detections. We need to apply a multi-channel-based decision mechanism in order to reduce the false detections, similar to the Wiener filter-based method suggested by Van Eyndhoven *et al.*[39]

5. Conclusions and Future Work

We have developed an efficient, automated CNN-based IED detector with linear complexity to the number of channels. The proposed IED detector was evaluated on a sizable dataset of 554 subjects from MGH, and it achieved a false positive rate of 0.2 per minute for a sensitivity of 80%. The CNN IED was proved to be noninferior to neurologists by evaluating it on the MUSC dataset. Further, the IED detector was clinically validated at NUH with an accuracy of 81.41% in agreement with a single clinical expert. This system could ultimately aid neurologists in the diagnosis of epilepsy.

Deep learning systems have been shown to improve their performance with an increase in the size of the dataset. In the future, more annotated EEG data could be eventually added to train the system, thereby adaptively improving the performance of the detector. Detection of only a few unambiguous IEDs is necessary to conclude that an EEG is epileptic, from a clinical perspective. In the future, we plan to design a system to classify entire EEGs as IED-free versus IED-present, leveraging on the IED detector proposed in this study. Such an EEG classification system may ultimately provide efficient automated monitoring of EEG recordings for the diagnosis of epilepsy. However, this could be challenging as the sensitivity of EEG classification based on IEDs is low as well the IED occurrences are highly subjective.[7,10] Further, IED frequency has shown to exhibit certain rhythmicity with the occurrence of seizures.[40] Certain seizure patterns are observed to contain IED patterns.[5] Moreover, the IED location is relevant in order to identify the localization of epileptogenic foci. Consequently, the proposed system could be useful for seizure detection and prediction, as well as aiding in epilepsy surgery.

Acknowledgments

The data collection at MGH was performed under the supervision of Dr. Westover, Dr. Jing, and Dr. Cash. The data collection at MUSC was performed under the supervision of Dr. Halford. We acknowledge the support of the Critical Care EEG Monitoring Research Consortium for the support during the MUSC dataset annotation. The clinical validation at NUH, Singapore, was performed under the supervision of Dr. Rahul Rathakrishnan, and it was supported by the National Health Innovation Centre (NHIC) Grant (NHIC-I2D-1608138).

Appendix A.: IED Detection Literature

Table A.1.

Overview of IED detection studies.

Reference	EEG count IED count EEG channels	Algorithm	Performance measures			
			Sensitivity	Specificity	Precision	FP/min
Current study	461 + 200 + 75	1D CNN	80%	—	82%	0.2
	18,164 + 235 + 136	Ensemble CART	80%	—	30.3%	2.55
	Single-channel	SVM	80%	—	29.5%	4.72

Reference	EEG count IED count EEG channels	Algorithm	Performance measures			
			Sensitivity	**Specificity**	**Precision**	**FP/min**
Jing *et al.*[5] (2019)	9571	2D CNN		AUC = 0.980		
	13,262	Persyst 13		AUC = 0.882		
	19-channel					
Clarke *et al.*[22](2019)	103 + 7					
	6000 + 6983					
	21-Channel	2D CNN	96.7%	—	—	1.16
Thomas *et al.*[4] (2018)	156					
	18,164					
	Single-channel	1D CNN	80%	—	55%	2.38
	100					
	1815			98% (epileptic)		0.6 (epileptic)
Tjepkema-Cloostermans *et al.*[21] (2018)	19-Channel	2D CNN	47.4%	99% (nonepileptic)	—	0.03 (nonepileptic)
					—	
Le *et al.*[11] (2018)	19					
	1491					
	Single-channel	DBN	87.35%	97.89%	—	—
Fukami *et al.*[12] (2018)	8	Eigenvalue				
	469	analysis				
	19-channel	Clustering	70.8%	—	39%	—
Ganglberger *et al.*[37] (2017)	12					
	5582	EpiSpike 1.5	34.6%	—	58.5%	0.33
	Multi-channel	ANN	44.1%	—	56.2%	0.33
Thomas *et al.*[38] (2017)	50					
	8929					
	Single-channel	AP–TM	90%	85.9%	66.8%	—
Le Douget *et al.*[41] (2017)	17					
	2157					
	Single-channel	WT, RF	62%	—	26%	—
Scheuer *et al.*[36] (2017)	40					
	5474	Human experts	51.5%			1.99
	Multi-channel	Persyst 13	43.9%	—	—	1.65
Johansen *et al.*[20] (2016)	5					

Reference	EEG count IED count EEG channels	Algorithm	Performance measures			
			Sensitivity	Specificity	Precision	FP/min
	7500					
	Single-channel	CNN		AUC = 0.947		
Carey *et al.*[42] (2016)	6					
	NR					
	Single-channel	NN	82.68%	—	72.67%	—
Rosado and Rosa[43] (2016)	NR + 3					
	NR + 1308	WT				
	Single-channel	Fuzzy Logic	80%	70%	—	—
Zacharaki *et al.*[13] (2016)	1					
	101					
	Multi-channel	SVM	97%	63%	—	0.1
Lodder and van Putten[15] (2014)	8 + 15					
	482 + 241	Self-adapting				
	Single-channel	TM	92%	—	—	6
Lodder *et al.*[14] (2013)	8 + 15					
	482 + 241					
	Single-channel	TM	90%	—	—	2.36
Liu *et al.*[44] (2013)	15					
	142	AdaBoost				
	Single-channel	classifier	95.5%	92.4%	—	—
Halford *et al.*[31] (2013)	100					
	2571	ANN				
	Single-channel	WT	65.3%	80.6%	—	—

Notes: CNN: Convolutional Neural Network, DBN: Deep Belief Network, ANN: Artificial Neural Network, AP: Affinity Propagation, WT: Wavelet Transform, SVM: Support Vector Machine, AUC: area under the receiver operating characteristic curve, RF: Random Forest, TM: Template Matching, NR: Not Reported, NA: Not Available and "+": Multiple sets and CART: Classification and Regression Tree.

Appendix B.: Different Parameters of the Traditional Classifiers Optimized

Table B.1.

Different parameters of the traditional classifiers optimized.

Classifier	Parameters	Range/Type
CART	Minimum number of leaf node observations	$\left[1, \max\left(2, \left[\frac{N}{2}\right]\right)\right]$
	Maximal number of decision splits	$[1, \max(2, N-1)]$
	Split criterion	Gini's diversity index, Cross-entropy
k-NN	Distance measure	Euclidean, Minkowski
	Number of neighbors *k*	[3, 100]
LDA	Distance measure	Euclidean, Minkowski
	Discriminant type	Linear
	Delta	$[10^{-6}, 10^{3}]$
	Gamma	[0, 1]
Ensemble	Aggregation method	AdaBoost
	Weak learners	*k*-NN, LDA, CART
	Number of weak learners	[10, 500]
	Learning rate	$[10^{-3}, 1]$
SVM	Kernel	Linear, Gaussian
	Box constraint	$[10^{-3}, 10^{3}]$
	Kernel scale parameter	$[10^{-3}, 10^{3}]$

Appendix C.: Performance Measures

We evaluate the IED detectors based on the following performance measures:

- **True positives (TP).** The number of true IEDs predicted as IEDs.
- **False negatives (FNs).** The number of true IEDs predicted as non-IEDs.
- **True negatives (TNs).** The number of true non-IEDs predicted as non-IEDs.
- **False positives.** The number of non-IEDs predicted as IEDs.
- **Sensitivity/Recall.** The proportion of IEDs that are correctly identified as IEDs from a set of IED and non-IED waveforms:

$$\text{Sensitivity} = \frac{\text{TPs}}{\text{TPs} + \text{FNs}}. \tag{C.1}$$

- **Specificity.** The proportion of non-IEDs or background waveforms that are correctly classified as non-IEDs:

$$\text{Specificity} = \frac{\text{TNs}}{\text{TNs} + \text{FPs}}. \tag{C.2}$$

- **Precision/Selectivity.** The proportion of true IEDs in the entire set of waveforms that are predicted as IEDs:

$$\text{Precision} = \frac{\text{TPs}}{\text{TPs} + \text{FPs}}. \tag{C.3}$$

- **Accuracy.** The ratio of correct predictions to total predictions made:

$$\text{Accuracy} = \frac{\text{TPs} + \text{TNs}}{\text{TPs} + \text{TNs} + \text{FPs} + \text{FNs}}. \tag{C.4}$$

- ***F*1-score.** The harmonic mean of precision and recall:

$$F1\text{-score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}. \tag{C.5}$$

- **False positives per minute (false detections reported per minute or hour).** It is the rate at which non-IEDs are misclassified as IEDs.
- **Receiver operating characteristic (ROC) curve.** A graphical plot between true positive rate (sensitivity) and false positive rate (1−specificity) for various thresholds.
- **Precision–recall curve.** A graphical plot between Precision and Recall for various thresholds.
- **Area under the ROC curve.** The area under the ROC curve is referred to as AUC. It gives the probability that the system labels a true IED over a randomly chosen non-IED as an IED. A high AUC value indicates a better system performance.
- **Area under the precision–recall curve.** The area under the precision–recall curve is referred to as AUPRC.

References

1. Thijs RD, Surges R, O Brien TJ and Sander JW, Epilepsy in adults, The Lancet 393 (2019) 689–701.
2. Epilepsy Foundation of America, About Epilepsy: The basics (2019), http://epilepsy.com/Starthere/about-epilepsy-basics.
3. Halford JJ, Computerized epileptiform transient detection in the scalp electroencephalogram: Obstacles to progress and the example of computerized ecg interpretation, Clin. Neurophysiol 120(11) (2009) 1909–1915. [PubMed: 19836303]
4. Thomas J, Comoretto L, Jin J, Dauwels J, Cash SS and Westover MB, EEG classification via convolutional neural network-based interictal epileptiform event detection, in Proc. 2018 40th Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC) (IEEE, 2018), pp. 3148–3151.
5. Jing J et al., Development of expert-level automated detection of epileptiform discharges during electroencephalogram interpretation, JAMA Neurology (2019), doi:10.1001/Jamaneurol.2019.3485.
6. Shanir PM, Khan KA, Khan YU, Farooq O and Adeli H, Automatic seizure detection based on morphological features using one-dimensional local binary pattern on long-term EEG, Clin. EEG Neurosci 49(5) (2018) 351–362. [PubMed: 29214865]

7. Smith S, Eeg in the diagnosis, classification, and management of patients with epilepsy, J. Neurol. Neurosurg. Psychiatry 76(Suppl 2) (2005) ii2–ii7. [PubMed: 15961864]
8. Celesia GG and Chen R-C, Parameters of spikes in human epilepsy, Dis. Nerv. Syst 37(5) (1976) 277–281. [PubMed: 944121]
9. Deuschl G et al., Recommendations for the practice of clinical neurophysiology: Guidelines of the International Federation of Clinical Neurophysiology, Electroencephalogr. Clin. Neurophysiol. Suppl 52 (1999) 1–304. [PubMed: 10617380]
10. Pillai J and Sperling MR, Interictal EEG and the diagnosis of epilepsy, Epilepsia 47 (2006) 14–22.
11. Le TX, Le TT, Dinh VV, Tran QL, Nguyen LT and Nguyen DT, Deep learning for epileptic spike detection, VNU J. Sci., Comput. Sci. Commun. Eng 33(2) (2018) 1–13.
12. Fukami T, Shimada T and Ishikawa B, Fast EEG spike detection via eigenvalue analysis and clustering of spatial amplitude distribution, J. Neural Eng 15(3) (2018) 036030. [PubMed: 29560928]
13. Zacharaki EI, Mporas I, Garganis K and Megalooikonomou V, Spike pattern recognition by supervised classification in low dimensional embedding space, Brain Inform 3(2) (2016) 73–83. [PubMed: 27747608]
14. Lodder SS, Askamp J and van Putten MJ, Inter-ictal spike detection using a database of smart templates, Clin. Neurophysiol 124(12) (2013) 2328–2335. [PubMed: 23791532]
15. Lodder SS and van Putten MJ, A self-adapting system for the automated detection of inter-ictal epileptiform discharges, PLos ONE 9(1) (2014) e85180. [PubMed: 24454813]
16. Krizhevsky A, Sutskever I and Hinton GE, ImageNet classification with deep convolutional neural networks, in Advances in Neural Information Processing Systems 25 (Curran Associates, Inc., 2012), pp. 1097–1105.
17. Ansari AH, Cherian PJ, Caicedo A, Naulaers G, De Vos M and Van Huffel S, Neonatal seizure detection using deep convolutional neural networks, Int. J. Neural Syst 29(4) (2019) 1850011. [PubMed: 29747532]
18. Acharya UR, Oh SL, Hagiwara Y, Tan JH and Adeli H, Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals, Comput. Biol. Med. 100 (2018) 270–278. [PubMed: 28974302]
19. Lin L-C, Ouyang C-S, Wu R-C, Yang R-C and Chiang C-T, Alternative diagnosis of epilepsy in children without epileptiform discharges using deep convolutional neural networks, Int. J. Neural Syst. 8 (2019) 1850060.
20. Johansen AR, Jin J, Maszczyk T, Dauwels J, Cash SS and Westover MB, Epileptiform spike detection via convolutional neural networks, in Proc. 2016 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP) (IEEE, 2016), pp. 754–758.
21. Tjepkema-Cloostermans MC, de Carvalho RC and van Putten MJ, Deep learning for detection of focal epileptiform discharges from scalp EEG recordings, Clin. Neurophysiol. 129 (2018) 2191–2196. [PubMed: 30025804]
22. Clarke S et al., Computer-assisted EEG diagnostic review for idiopathic generalized epilepsy, preprint (2019), bioRxiv, 682112, doi:10.1101/682112.
23. Peterson LE, *K*-nearest neighbor, Scholarpedia 4(2) (2009) 1883.
24. Breiman L, Friedman J, Stone CJ and Olshen RA, Classification and Regression Trees (CRC Press, 1984).
25. Hearst MA, Dumais ST, Osuna E, Platt J and Scholkopf B, Support vector machines, IEEE Intell. Syst. Appl 13(4) (1998) 18–28.
26. Mika S, Ratsch G, Weston J, Scholkopf B and Mullers K-R, Fisher discriminant analysis with kernels, in Proc. 1999 IEEE Signal Processing Society Workshop: Neural Networks for Signal Processing IX (IEEE, 1999), pp. 41–48.
27. Zhu J, Zou H, Rosset S and Hastie T, Multi-class AdaBoost, Stat. Interface 2(3) (2009) 349–360.
28. Jing J, Dauwels J, Rakthanmanon T, Keogh E, Cash S and Westover M, Rapid annotation of interictal epileptiform discharges via template matching under dynamic time warping, J. Neurosci. Methods 274 (2016) 179–190. [PubMed: 26944098]

29. Halford JJ et al., Interictal epileptiform discharge detection in EEG in different practice settings, J. Clin. Neurophysiol 35(5) (2018) 375–380. [PubMed: 30028830]
30. Halford JJ et al., Characteristics of EEG interpreters associated with higher interrater agreement, J. Clin. Neurophysiol 34(2) (2017) 168–173. [PubMed: 27662336]
31. Halford JJ et al., Standardized database development for EEG epileptiform transient detection: Eegnet scoring system and machine learning analysis, J. Neurosci. Methods 212(2) (2013) 308–316. [PubMed: 23174094]
32. Zaidel A, Spivak A, Grieb B, Bergman H and Israel Z, Subthalamic span of β oscillations predicts deep brain stimulation efficacy for patients with Parkinson's disease, Brain 133(7) (2010) 2007–2021. [PubMed: 20534648]
33. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow I, Harp A, Irving G, Isard M, Jia Y, Jozefowicz R, Kaiser L, Kudlur M, Levenberg J, Mané D, Monga R, Moore S, Murray D, Olah C, Schuster M, Shlens J, Steiner B, Sutskever I, Talwar K, Tucker P, Vanhoucke V, Vasudevan V, Viégas F, Vinyals O, Warden P, Wattenberg M, Wicke M, Yu Y and Zheng X, TensorFlow: Large-scale machine learning on heterogeneous systems, preprint (2016), arXiv:1603.04467 [cs.DC].
34. Martinez-Cantin R, Bayesopt: A bayesian optimization library for nonlinear optimization, experimental design and bandits, J. Mach. Learn. Res 15(1) (2014) 3735–3739.
35. Gotman J, Quantitative measurements of epileptic spike morphology in the human EEG, Electroencephalogr. Clin. Neurophysiol 48(5) (1980) 551–557. [PubMed: 6153963]
36. Scheuer ML, Bagic A and Wilson SB, Spike detection: Inter-reader agreement and a statistical turing test on a large data set, Clin. Neurophysiol 128(1) (2017) 243–250. [PubMed: 27913148]
37. Ganglberger W, Gritsch G, Hartmann MM,Fürbass F, Perko H, Skupch A and Kluge T, A comparison of rule-based and machine learning methods for classification of spikes in EEG, J. Commun 12(10) (2017) 589–595.
38. Thomas J, Jin J, Dauwels J, Cash SS and Westover MB, Automated epileptiform spike detection via affinity propagation-based template matching, in Proc. 2017 39th Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC) (IEEE, 2017), pp. 3057–3060.
39. Van Eyndhoven S, Hunyadi B, Dupont P, Van Paesschen W and Van Huffel S, Semiautomated EEG enhancement improves localization of ictal onset zone with EEG-correlated FMRI, Front. Neurol 10 (2019) 805. [PubMed: 31428036]
40. Karoly PJ, Freestone DR, Boston R, Grayden DB, Himes D, Leyde K, Seneviratne U, Berkovic S, O?Brien T and Cook MJ, Interictal spikes and epileptic seizures: Their relationship and underlying rhythmicity, Brain 139 (2016) 1066–1078. [PubMed: 26912639]
41. Le Douget J, Fouad A, Filali MM, Pyrzowski J and Le Van Quyen M, Surface and intracranial EEG spike detection based on discrete wavelet decomposition and random forest classification, in Proc. 2017 39th Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC) (IEEE, 2017), pp. 475–478.
42. Carey HJ, Manic M and Arsenovic P, Epileptic spike detection with EEG using artificial neural networks, Human System Interactions (HSI), 2016 9th Int. Conf., (IEEE 2016), pp. 89–95.
43. Rosado A and Rosa AC, Automatic detection of epileptiform discharges in the EEG, preprint (2016), arXiv:1605.06708 [cs.CV].
44. Liu Y-C, Lin C-CK, Tsai J-J and Sun Y-N, Model-based spike detection of epileptic EEG data, Sensors 13(9) (2013) 12536–12547. [PubMed: 24048343]

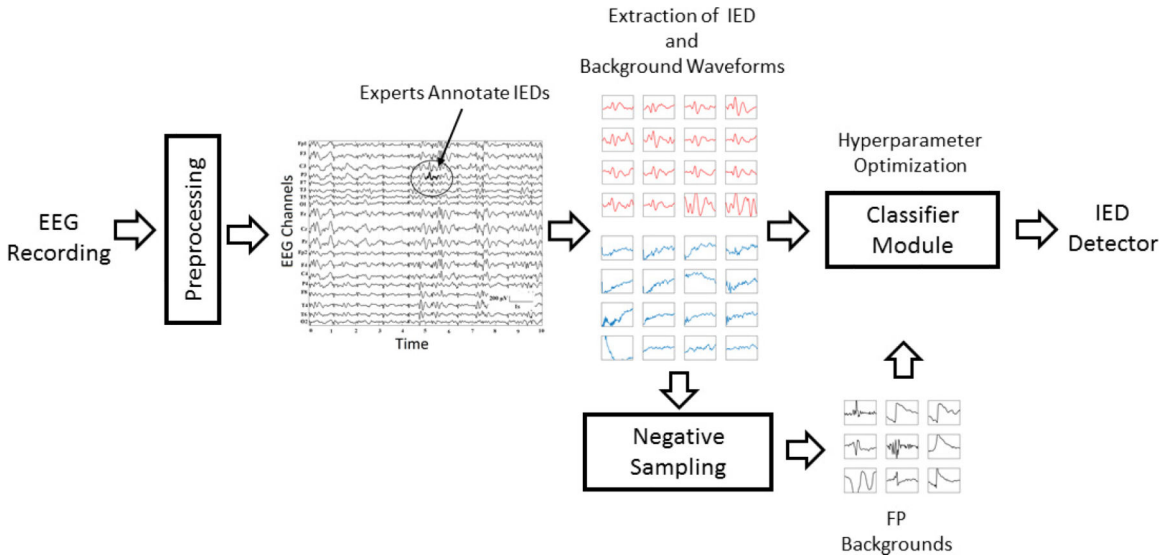


Fig. 1.
Block diagram of the proposed IED detector training scheme.

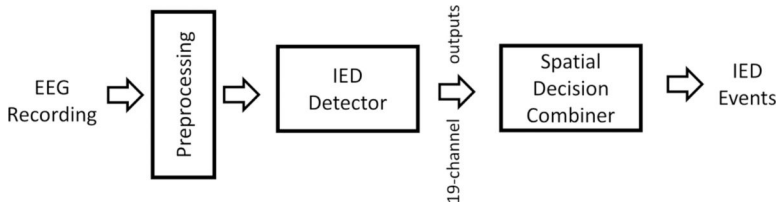


Fig. 2.
Block diagram of the proposed IED detector evaluation scheme.

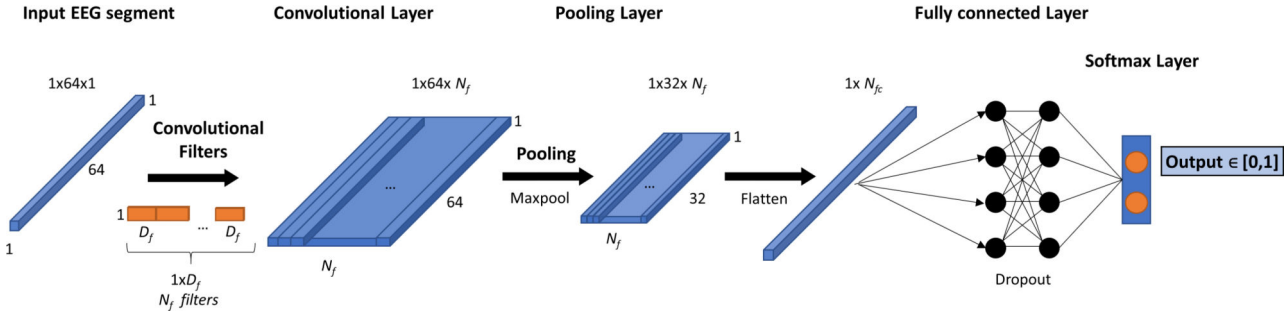


Fig. 3.

The proposed CNN architecture with single convolutional, pooling, and fully connected layers applied in this study. N_f indicates the number of convolutional filters, D_f the dimension of convolutional filters, and N_{fc} the number of neurons in the fully connected layer.

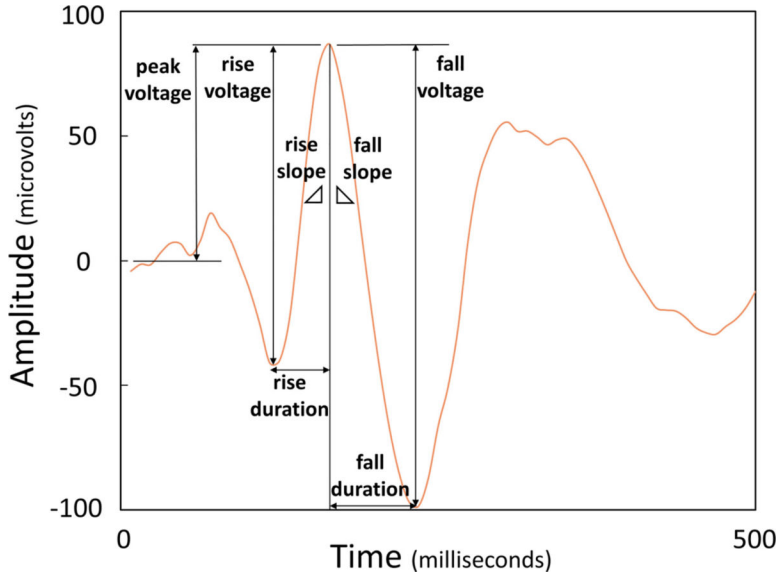
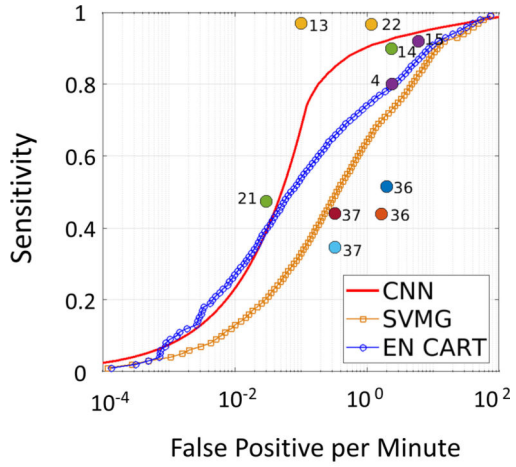
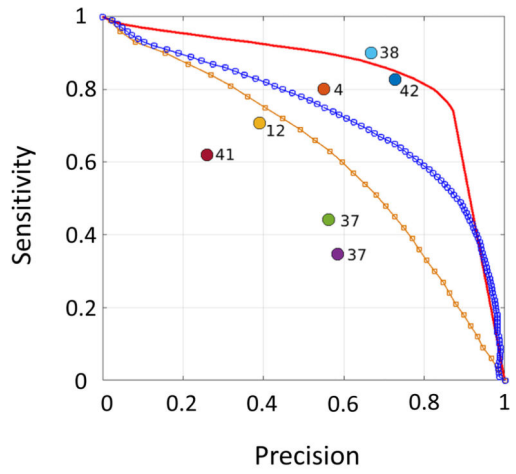


Fig. 4.

Morphological features of IEDs evaluated in this study. The features are amplitude (peak, rise, and fall amplitudes), duration (rise and fall durations), and slope (rise and fall slopes).



(a)



(b)

Fig. 5.

Performance comparison of the literature with the proposed CNN IED detector and traditional classifier-based IED detectors: (a) sensitivity–false positive rate curves and (b) sensitivity-precision curves. The performances of different studies in the literature are illustrated as circles.

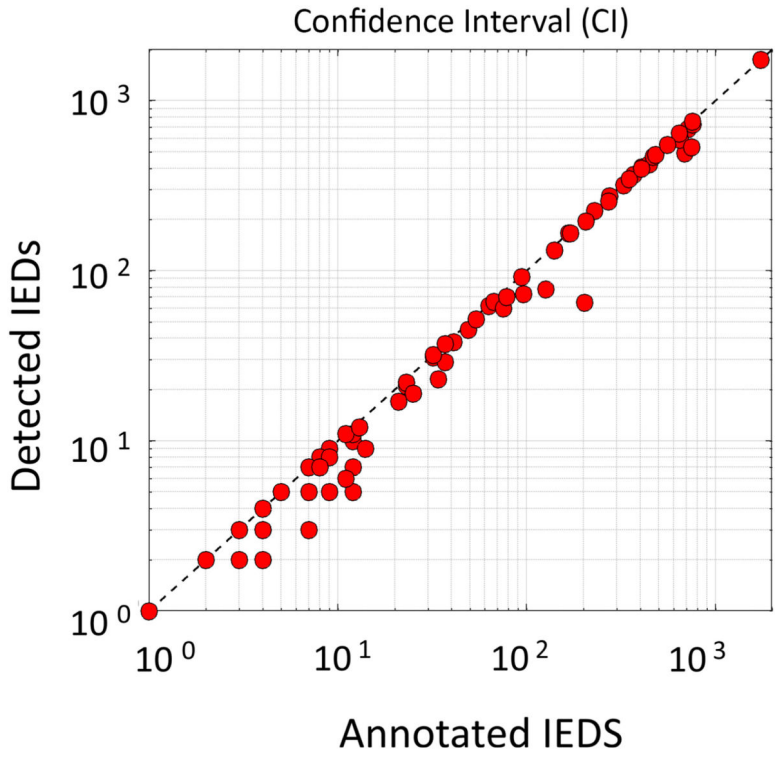


Fig. 6.

Comparison of the number of IEDs detected by the CNN IED detector versus the number of IEDs annotated by the experts for the 93 EEGs. The detections by the system are in agreement with the annotators for most of the EEGs.

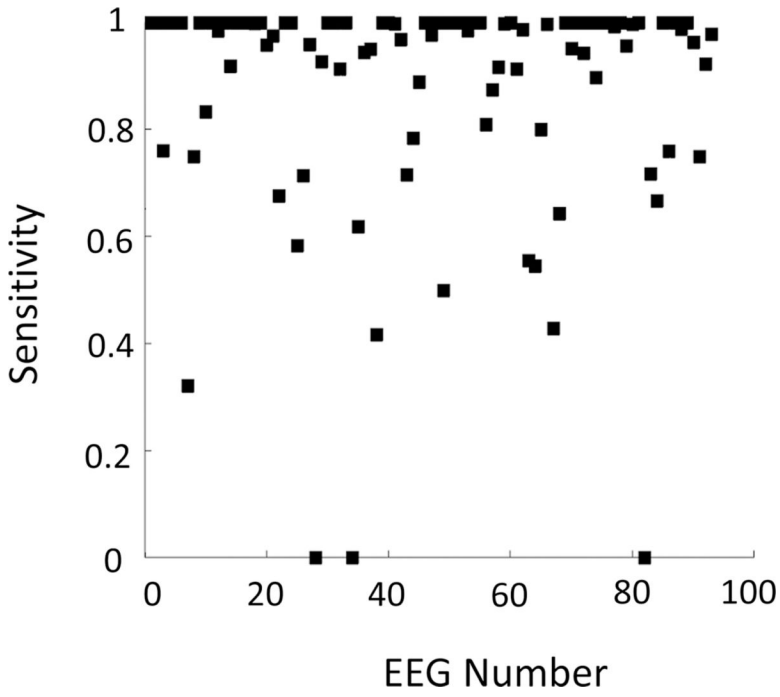


Fig. 7.
IED detection sensitivity for the 93 EEGs. The sensitivity is high for most of the EEGs.

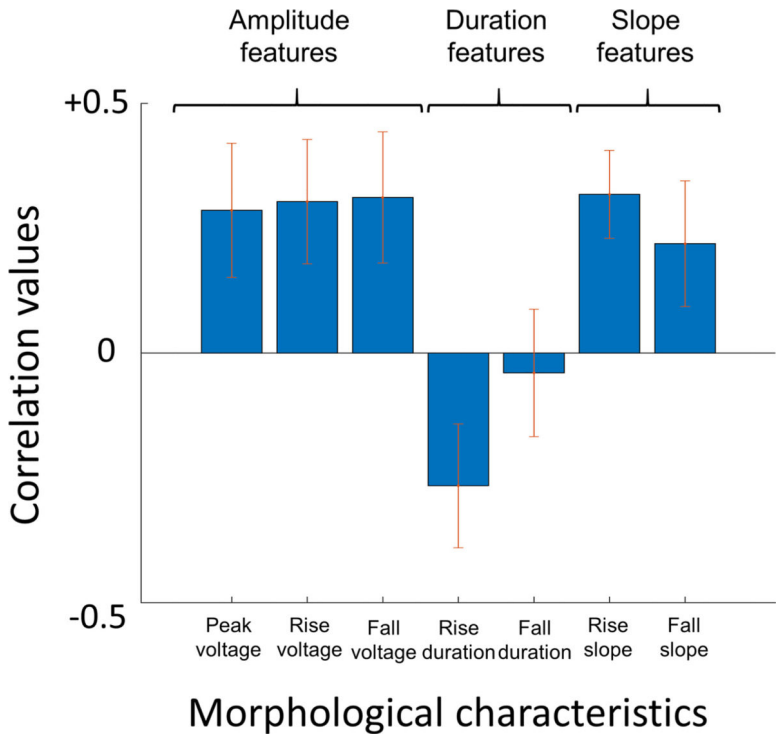


Fig. 8.
Correlation values of the different IED characteristics with the CNN output values.

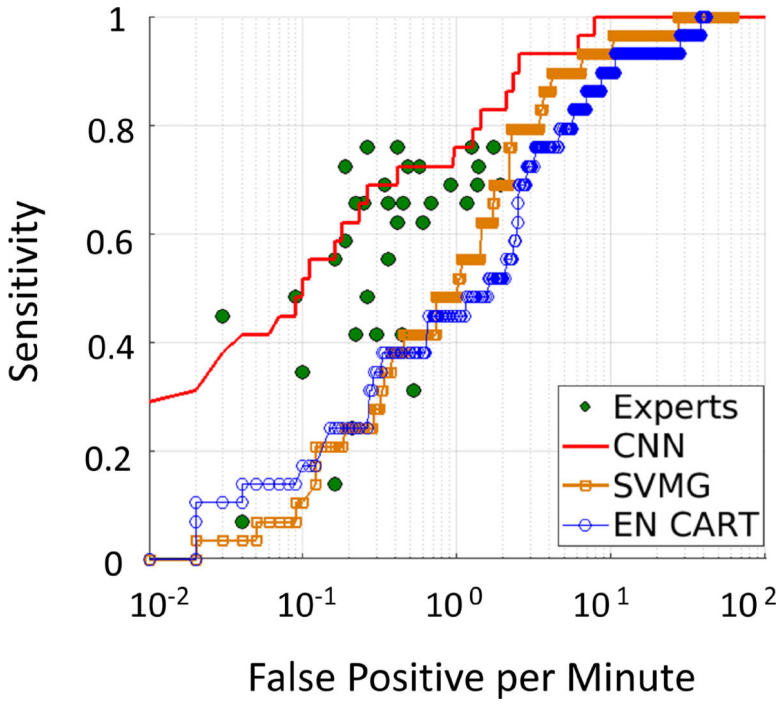


Fig. 9.

(Color online) Comparison of the proposed CNN IED detector and standard IED detectors (SVMG and Ensemble CART) based on the precision–sensitivity curve on the MUSC dataset with a mean annotator consensus threshold of 4. The performance of experts is illustrated as circles (green).

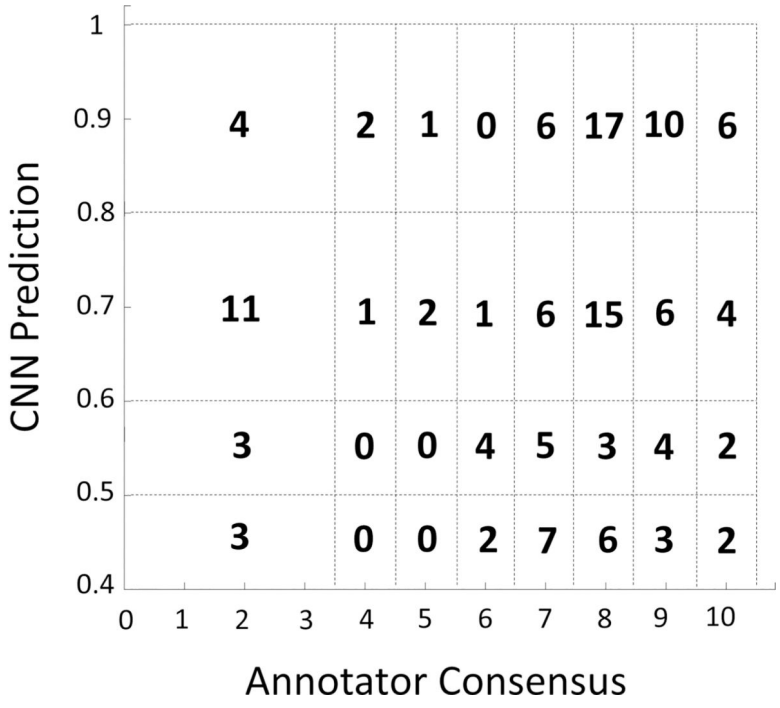


Fig. 10.

Comparison between the CNN IED detector predictions with the scores of expert (Rahul Rathkrishnan). There is a significant agreement between the system predictions and the annotator consensus.

Table 1.

Age distribution for the MGH EEG dataset (male/female).

Patient age range	Routine EEG	EMU	ICU	Total
0 to <5	2 (0/2)	3 (1/2)	—	5 (1/4)
5 to <18	22 (14/8)	6 (4/2)	1 (0/1)	29 (18/11)
18 to <30	6 (2/4)	4 (0/4)	—	10 (2/8)
30 to <50	5 (3/2)	3 (2/1)	—	8 (5/3)
50 to <70	18 (12/6)	1 (1/0)	—	19 (13/6)
>70	11 (4/7)	2 (0/2)	—	13 (4/7)
Total	64 (35/29)	19 (8/11)	1 (0/1)	84 (43/41)

Note: Routine EEG (outpatient and inpatient services). EMU: Epilepsy Monitoring Unit and ICU: Intensive Care Unit.

Table 2.

Age distribution for the NUH EEG dataset (male/female).

Patient age range	Routine EEG	ICU	Total
18 to <30	5 (2/3)	6 (6/0)	11 (8/3)
30 to <50	4 (2/2)	6 (4/2)	10 (6/4)
50 to <70	11 (5/6)	17 (9/8)	28 (14/14)
>70	16 (8/8)	10 (7/3)	26 (15/11)
Total	36 (17/19)	39 (26/13)	75 (43/32)

Note: Routine EEG (outpatient and inpatient services). ICU: Intensive Care Unit.

Table 3.

Distribution of the MGH dataset over the different cross-validation folds.

Fold number	Number of epileptic EEGs	Number of IED events	Number of annotated IEDs	Number of nonepileptic EEGs	Number of background event segments ($\times 10^6$)
1	18	2920	4077	93	1.34
2	19	2757	3571	92	1.47
3	18	2831	3207	92	1.47
4	19	2781	4021	92	1.37
5	19	2881	3288	92	1.23
Total	93	14,170	18,164	461	6.88

Table 4.

Different parameter values evaluated for CNN optimization.

Parameter	Values/Types
Number of convolution layers	1–4
Number of pooling layers	1–3
Number of fully connected layers	1–3
Number of convolution filters	4, 8, 16, 32, 64
Dimension of convolution filters	1 × 3, 1 × 4, 1 × 5, 1 × 6, 1 × 7, 1 × 8
Number of hidden layer neurons	100, 500, 1000
Activation functions	ReLu, tanh, sigmoid
Dropout probability	0.5
Size of the batch processing	$2n_s, n_s, \frac{n_s}{2}, \frac{n_s}{4}$
Maximum number of iterations	20,000
Optimizer	Adam
Learning rate	10^{-4}
Measure	Cross-entropy

Note: n_S: Number of IEDs.

Table 5.

The five-fold cross-validation AUC and AUPRC values for the different IED detectors.

Method	AUC	AUPRC
SVMG	0.988 ± 0.01	**0.615 ± 0.19**
SVML	0.954 ± 0.04	0.177 ± 0.10
CART	0.856 ± 0.03	0.376 ± 0.08
k-NN	0.925 ± 0.03	0.489 ± 0.01
LDA	0.973 ± 0.01	0.159 ± 0.07
Ensemble CART	0.991 ± 0.00	**0.730 ± 0.10**
k-NN	0.908 ± 0.03	0.121 ± 0.01
LDA	0.971 ± 0.01	0.431 ± 0.13
CNN	0.989 ± 0.01	**0.838 ± 0.04**

Note: The data is reported as mean ± standard deviation.

Table 6.

The five-fold cross-validation false positive rate comparison for fixed sensitivity thresholds for the selected IED detectors.

Method	FP rate (per minute) for sensitivity = 80% (*F*1-score)	FP rate (per minute) for sensitivity = 90% (*F*1-score)
SVMG	4.72 ± 5.43	11.57 ± 12.30
	0.433	0.242
Ensemble CART	2.55 ± 3.48	8.61 ± 10.68
	0.439	0.218
CNN	**0.2 ± 0.11**	**0.92 ± 0.83**
	0.809	**0.682**

Note: The data is reported as mean ± standard deviation.