Seizure Prediction in Scalp EEG Using 3D Convolutional Neural Networks With an Image-Based Approach

Ahmet Remzi Ozcan and Sarp Erturk, *Senior Member, IEEE*

***Abstract*— Epileptic seizures occur as a result of a process that develops over time and space in epileptic networks. In this study, we aim at developing a generalizable method for patient-specific seizure prediction by evaluating the spatio-temporal correlation in the features obtained from multichannel EEG signals. Spectral band power, statistical moment and Hjorth parameters are used to reveal the frequency and time domain features of the EEG signals. The features are given as input to a convolutional neural network (CNN) by transforming into a sequence of multi-color images according to the topology of the EEG channels. The multi-frame 3D CNN model is proposed to evaluate the temporal and spatial correlation in training data collectively. The proposed 3D CNN model achieves a sensitivity of 85.7%, a false prediction rate of 0.096/h, and a proportion of time-in-warning of 10.5%, in the tests performed with 16 patients from the CHB-MIT scalp EEG dataset. The results show that the superiority of the proposed method to a Poisson based random predictor was statistically significant for 93.7% of the patients, at significance level of 0.05. Our experiments with various timing constraints show that epileptic stage lengths are an important factor affecting seizure performance. We present a subject-specific seizure prediction method that is robust for unbalanced data and can be generalized to any scalp EEG dataset without the need for subject-specific engineering.**

***Index Terms*— Convolutional neural networks, epilepsy, epileptic stage length, scalp EEG, seizure prediction, subject-specific modelling, unbalanced classification.**

I. Introduction

EPILEPSY is a common neurological disorder characterized by abrupt and recurrent seizures; affecting more than 40 million people worldwide. Epileptic seizures occur occasionally as a result of excessive and synchronous neuronal discharges. Improvements in the analysis of electroencephalographic (EEG) recordings which reveal information about the dynamics of the epileptic brain have led to an increase in EEG- based approaches to predict upcoming seizures and to localize the seizure onset zone. Studies on EEG data from patients admitted for epilepsy surgery showed strong evidence that seizures begin to develop minutes to hours before the clinical onset [1]. The majority of seizure prediction studies assume that there are four consecutive brain activity states in the epileptic EEG signals, namely interictal, preictal, ictal and postictal. However, some recent studies have adopted an approach that provides a continuous forecast of seizure probability rather than a categorical seizure prediction in which epileptic EEG signals are classified as distinct activity states [2]–[6]. In accordance this approach, the pro-ictal state is defined as a period of increased seizure risk [4].

Generally, the aim is to distinguish the interictal stage from the preictal stage, which corresponds to the situation just prior to the ictal stage where the seizure occurs clinically. This approach turns seizure prediction into a binary classification problem in which epileptic EEG samples are classified as either interictal or preictal. Although studies on clinic EEG data revealed that the extracted features begin to differentiate before the ictal stage [1], [7], the fact that most of the studies that use same measurements or features failed to provide acceptable performance for some patients in different datasets shows that the problem of epileptic seizure prediction remains challenging [8]–[11]. The timing, characteristics, and dynamics of the epileptic stages vary widely among the subjects, hence a general method of achieving high seizure prediction performance for each patient without being specifically trained for the concerned patient is yet to be presented. Therefore, in most of the recent studies that achieve relatively high seizure prediction sensitivity, a patient-specific approach is adopted [9]–[11], [12]–[15]. There are several recording types of EEG, depending on where the signals are collected. Intracranial EEG (iEEG) signals are gathered with electrodes placed directly on the open surface of the brain while scalp EEG signals are collected with scalp electrodes. Although some studies have demonstrated that iEEG is suitable for the use in a long-term, implanted seizure advisory system [16], [17], scalp EEG is more commonly used in clinical practice as an easily applicable method and is widely used in studies because of its ease of access to data sets.

In previous studies on seizure prediction, many linear and nonlinear features have been reported to detect changes in EEG before the onset of seizure. Absolute and relative spectral band power [8], [13], [18], [19],

Manuscript received May 7, 2019; revised August 7, 2019; accepted September 17, 2019. Date of publication September 25, 2019; date of current version November 6, 2019. *(Corresponding author: Ahmet Remzi Ozcan.)*

A. R. Ozcan is with the Department of Mechatronics Engineering, Bursa Technical University, 16310 Bursa, Turkey, and also with the Kocaeli University Laboratory of Image and Signal Processing (KULIS), 41380 Kocaeli, Turkey (e-mail: ahmet.ozcan@btu.edu.tr).

S. Erturk is with the Department of Electronics and Telecommunication Engineering, University of Kocaeli, 41380 Kocaeli, Turkey (e-mail:sarp@ieee.org).

Digital Object Identifier 10.1109/TNSRE.2019.2943707

autoregressive coefficients [20], accumulated energy [21], Hjorth parameters, statistical moments, and the spectral edge frequency [22] are used as linear features. Various non-linear features have also been used including the largest Lyapunov exponent [23], correlation dimension [24], dynamical similarity index [25], phase synchronization [15], [26] and optimum allocation sampling (OAS) [27]. More complex representations such as wavelet transform [10], Q-factor wavelet transform [28], [29], short-time Fourier transform [11], fractional Fourier transform [30] and FFT/DWT combined with InfoGain [31] have recently been used for epileptic signal classification. Support vector machines [8], [19], [20], [32], Bayesian classifier [12], [33], convolutional networks [10], [11] and other machine learning algorithms have been employed for the purpose of classifying the extracted features.

In this study, we aimed to present a model that can successfully predicts epileptic seizures. We also investigated the spatial and temporal dynamics of epileptic activity, which is important in seizure localization, using our training model. We considered that the most appropriate approach for this purpose is to evaluate the features obtained from the EEG signals as an image in which preserving the spatial structure. In preliminary study investigating the applicability of proposed approach, we performed a patient-independent seizure prediction with multi-color image series derived from multichannel EEG signals using cascade CNN and LSTM networks [34]. One of the major drawbacks in this study is that training and testing were performed with randomly selected data from all patients and using k-fold cross-validation instead of leave-one-out cross-validation. Another problem is that the spatial correlation was evaluated starting from the first layer in the CNN, while, the evaluation of temporal correlation was postponed until LSTM networks in the structure. In the present work, we improve on our previous study and propose a 3D CNN-based patient-specific seizure prediction method to evaluate both spatial and temporal relevance in multichannel EEG signals starting from the first layer. In addition we used leave-one-out cross validation to ensure that the performance of the test data was similar to the real conditions.

This paper is organized as follows. Section II first describes the utilized data set, used feature extraction methods, and an approach that is used to construct image-like 2D data from EEG time series; followed by the proposed learning models, training phase, and generation of critical regions' heatmaps. Section III presents experimental results of the proposed models for various epileptic stage lengths. Discussion of results and comparison with other seizure prediction approaches are provided in Section IV. Finally, the conclusion is presented in Section V.

II. Materials and Methods

A. Dataset

In this study, the proposed models are trained and tested on the CHB-MIT scalp EEG dataset [35], [36], collected at the Children's Hospital Boston. The dataset consists of 844-hour continuous EEG recordings of 23 pediatric patients.

TABLE I
Total Evaluated Interictal Data Duration (in Hour) by Distance Between Interictal and Ictal Stages

Patient	Interictal-Ictal Distance			No. of Seizures
	60 Minute	120 Minute	240 Minute	
Pt 1	28.6	22.6	14.4	7
Pt 2	30.2	28.0	26.0	3
Pt 3	30.4	29.4	27.4	6
Pt 5	29.8	22.5	14.4	5
Pt 7	61.8	57.3	51.3	3
Pt 9	61.0	56.0	48.6	4
Pt 10	36.9	31.6	24.2	7
Pt 13	22.3	20.0	15.1	5
Pt 14	16.4	10.7	4.7	6
Pt 16	9.8	7.6	5.6	5
Pt 17	16.1	15.1	11.8	3
Pt 18	29.9	29.0	26.4	5
Pt 19	27.0	27.0	27.0	3
Pt 20	20.4	20.0	19.1	6
Pt 21	28.3	26.4	23.4	4
Pt 23	17.2	16.2	14.2	5
Total	466.1	419.4	353.5	77

EEG signals were collected from 22 electrodes at 256 Hz sampling rate using the international 10-20 system with a bipolar montage. During these EEG recordings 182 seizures occurred [35]. We defined the 30-minute signal before the onset of seizure as the preictal period. The 10-minute signal after the end of the seizure was defined as the postictal period. The 1-minute interval between the preictal period and the seizure was considered an intervention time and excluded from the training data. At least one hour before the onset of seizures and one hour after the end of seizures were determined as interictal periods. In cases where more than one seizure occurs close to each other, the incoming seizure is not evaluated if there is not enough preictal data. The minimum preictal period length was determined as 15 minutes. We evaluated patients who had at least three seizures in the EEG records and had at least three hours of interictal period. This is due to the fact that less than three preictal and interictal periods cause an overfitting problem in training with cross-validation. Considering all these definitions and constraints, 77 seizures and 466 hours of interictal data are available for 16 patients.

B. Feature Extraction

We used spectral band power, statistical moment and Hjorth parameters as features. These features are calculated using moving window analysis for each EEG channel separately [15]. In moving window analysis, we used a window length of 4 seconds and a half overlap. This provides a prediction result of a seizure every 2 seconds.

In classical EEG analysis, the frequency domain of the EEG signal is typically separated into 5 spectral bands: δ (0.5-4 Hz), θ (4-8 Hz), α (8-13 Hz), β (13-30 Hz) and γ (up to 30 Hz). Considering that high frequency sub-bands are effective in epileptic seizure estimation, the γ band is split into four sub-bands [19] γ-1 (30-50 Hz), γ-2 (50-75 Hz), γ-3 (75-100 Hz), γ-4 (100-128 Hz). The frequency range 57-63 Hz and 117-123 Hz are not used in spectral power

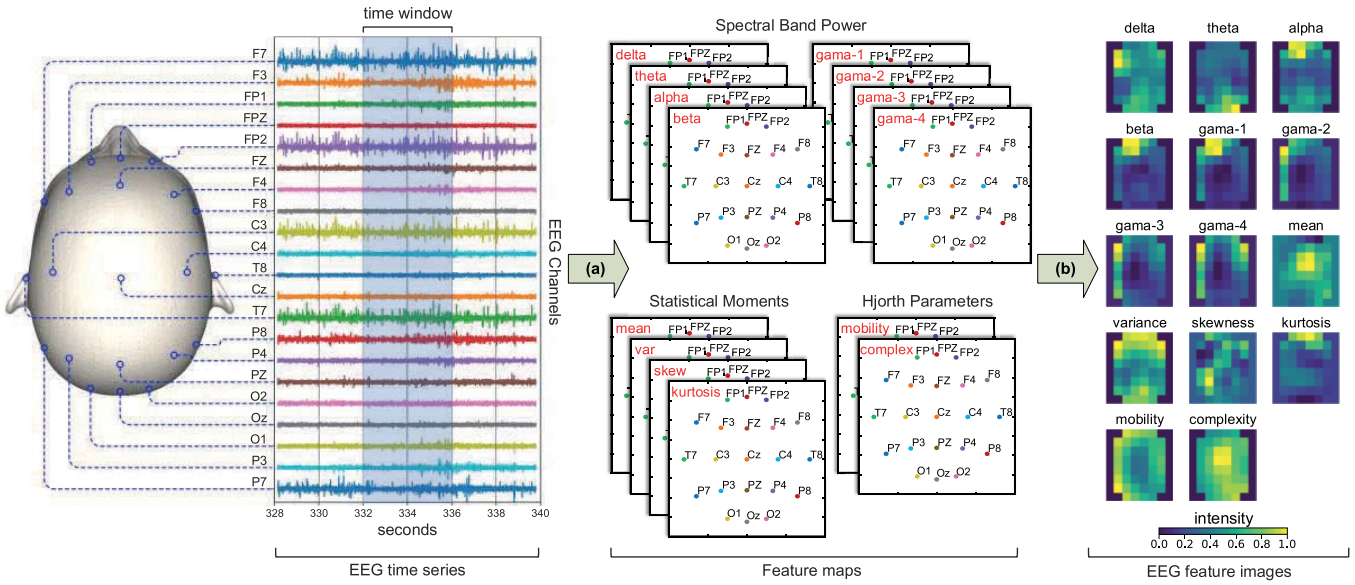


Fig. 1. Constructing EEG feature images from EEG time series. (a) Spectral band powers, statistical moments and Hjorth parameters are obtained using moving windows of 4 seconds and a half overlap from the EEG time series for each individual EEG channel. The features are placed on a 2D map obtained using the Azimuthal Unequal Projection [39] by the 3D distribution of the EEG channels. (b) For each map, an EEG feature image is constructed by piecewise cubic interpolation using a Clough-Tocher [40] scheme. The colors in the images represent the intensity of the pixel values.

calculation to eliminate the power line noise at 60 Hz and harmonics. The DC component at 0 Hz was also excluded.

Statistical moments give information about the amplitude distribution and shape of time series. It is stated that statistical moments may be useful for early seizure detection in high amplitude seizure signals [7]. In this study, as the first four statistical moments, mean, variance, skewness and kurtosis are used to extract time-domain features of EEG time series. Hjorth defines three time-domain parameters as indicators for the statistical properties of the signals [37]. These are activity, mobility, and complexity. A significant increase in the mobility and complexity of the EEG signal in the preictal stage is indicated [18]. We also used mobility and complexity parameters to extract the time-domain features of EEG signals.

C. Constructing Images From EEG Time Series

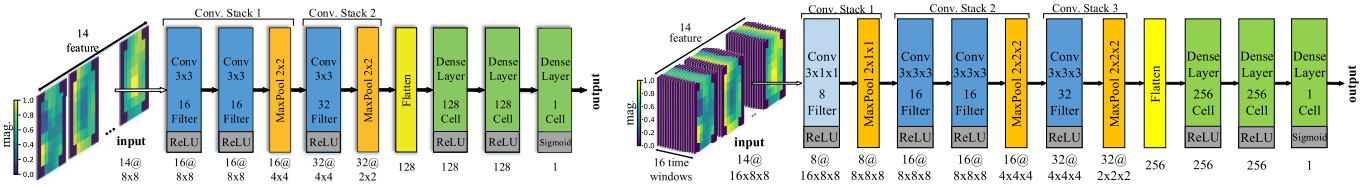
When creating a feature vector in EEG data analysis, evaluation of the data regardless of electrode placement causes the spatial and spectral relationship to be ignored. Therefore, to preserve the spatial structure in EEG data, it has been proposed to transform the features into a 2D image and use multiple color channels to represent the spectral dimension [38]. The EEG electrodes are placed in a 3D space on the scalp. In order to obtain 2-dimensional EEG image series from 3-dimensional EEG activity maps, it is necessary to project from 3D space onto a 2D surface. For this purpose, we used the Azimuthal Unequal Projection [39] to preserve the relative distance between adjacent electrodes during this transformation As a disadvantage in this method, the relative distances between all the electrode pairs are not exactly preserved, since the projection is intended to preserve the relative distance of each point with the reference point only.

EEG features projected on 2D surface are transformed to image format for use in the proposed CNN model. The pixel values of the image obtained in this transformation are calculated by piecewise cubic interpolation using a Clough-Tocher scheme [40]. The input image resolution for the proposed CNN model is selected as 8×8. The resulting image represents the spatial distribution of the extracted features over the cortex. Fig. 1 illustrates an overview of our approach to constructing EEG feature images from EEG time series.

D. Convolutional Neural Network

Convolutional neural networks, which are commonly used in computer vision applications, have also been used for epileptic seizure prediction [10], [11], [41]. We have followed an approach based on the VGGNet [42] in the design of CNN architectures. The VGGNet, a deep convolutional network developed by Oxford's Visual Geometry Group, consists of consecutive stacked convolutional and maximum pooling layers. The number of kernels within convolution layers is doubled in each new stack. Shallow layers in CNN reveal general features, while the deeper layers represent more detailed features. In most cases, adding more layers does not improve network performance. In fact, if redundant layer and filter are present in the network, this will cause overfitting problem in the training phase. Although it is still not clear how to design a deep CNN structure optimally, two criteria are defined as learning capacity and learning necessity for optimum design in convolutional neural networks [43]. These criteria transform the deep convolutional neural network design into a constrained optimization problem. The goal in this problem is to reach the maximum net depth within these two constraints.

1) Single-Frame CNN Model: Initially, a single-frame CNN model is proposed for 8×8 pixel single-frame input images



(a) The single-frame CNN model (b) The multi-frame 3D CNN model

Fig. 2. The single-frame and the multi-frame convolutional neural network architectures. (a) In the single-frame CNN model, 14 images of 8×8 pixel, derived from 14 features, are evaluated as input data. (b) In the multi-frame 3D CNN model, the sequence of images derived from consecutive windows is used as input to take into account the temporal evolution in brain activity, as well as the correlation among EEG channels. 14 images of 8×8 pixels in the past 16 time windows are evaluated at once. The colors in the images represent the intensity of the pixel values.

in accordance with the learning capacity and learning requirement criteria. The CNN model consists of two convolutional stacks and three fully-connected layers, as shown in Fig. 2(a). The first stack contains two convolutions and one max-pooling layer, while the second stack has one convolution and one max-pooling layer. In the first stack, the number of filters (convolution kernels) was determined to be 16, which is doubled in the second stack. The convolution kernel size is taken as 3×3 in all convolution layers with stride of 1×1. Image sizes are reduced by half by using 2×2 windows in max-pooling layers. The first two fully-connected layers consist of 128 cells while the last contains 1 cell as the output layer. In order to prevent the CNN model from overfitting during the training phase, both of the first two fully-connected layers have a 50% dropout rate. The rectified linear unit activation function is used for all convolutional and fully-connected layers except the output layer. With the use of the sigmoid activation function in the output layer, a classification result from 0 to 1 is obtained.

2) Multi-Frame CNN Model: Studies show that seizures occur as a result of temporal and spatial evolving processes in an epileptic network [1]. The single-frame CNN model operates in the spatial domain, so only the correlation between EEG channels is evaluated. A multi-frame 3D CNN model is proposed to evaluate the correlation among both time windows and EEG channels. The sequence of images derived from 16 consecutive time windows is used as input data in the 3D CNN model to account the temporal evolutions in brain activity.

In contrast to the CNN model, the 3D CNN model has three convolutional stacks followed by three dense layers, as shown in Fig. 2(b). The first convolutional stack contains one convolution and one max-pooling layer. The other stacks are the same as the CNN model. In the first stack, only the temporal correlation is evaluated by taking the convolution kernel size as 3×1×1. There are 8 filters in the layer. The image size is reduced to 8×8×8 by using 2×1×1 window in max-pooling layer. In other stacks, convolution kernel and max-pooling window sizes are determined as 3×3×3 and 2×2×2, respectively. In the 3D CNN model, the number of cells in the first two fully-connected layers is increased to 256.

E. Training

In this study, the training of the models is carried out separately for each subject. Convergence of the model parameters during training is achieved by optimizing the cross-entropy loss function, as in (1), where the true probability $p(x)$ is the true label, and the estimated probability $q(x)$ is the predicted value for sample x in the training set D. Since the output of the created network is binary, the cross-entropy loss function transforms into the form as in (2), where n is the number of samples, $y = [y_1, y_2, , y_n] \in \mathbb{R}^n$ are the output values of the network, $t = [t_1, t_2, , t_n] \in \mathbb{R}^n$ are the target values.

$$Cross - Entropy(p, q) = \sum_{x \in D} p(x) \log(q(x)) \quad (1)$$

$$E_n = -\frac{1}{n} \sum_{k=1}^{n} t_k . \log(y_k) + (1 - t_k) \log(1 - y_k) \quad (2)$$

We used leave-one-out cross-validation to evaluate the performance of the proposed approaches for each patient. If a patient data set contains n seizures and t-hour interictal recordings, the entire of interictal recordings are partitioned into n blocks. Each of these interictal blocks is approximately t/n hours and is randomly grouped with any of the preictal recordings. There are n runs for the patient, and in each run one of these interictal-preictal pairs is left for testing, while the remaining n-1 pairs are used during the training phase. The n-1 pair used in training phase is split into 5 folds according to the leave-one-out cross-validation method and 80% of randomly selected data is used in training, while the remaining 20% is used for validation to prevent overfitting. At the end of each training fold, the model is tested with the interictal - preictal pair reserved for testing. Then, the averages of obtained values with standard deviations are reported as the prediction results of the test data.

Although more iteration on the training phase increases the accuracy of training, it casues overfitting the training data. To prevent the CNN model from overfitting, we used early-stopping. In this method, the training is halted when error on the validation set begins to increase. In detail, network parameters on the past iterations are stored, and verification errors are monitored. When it is detected that the validation error has started to increase, the network parameters are restored to the values at the iteration where the lowest validation error is obtained.

Learning algorithms are designed in accordance with balanced class distributions or equal misclassification costs. The described epileptic stage lengths cause an imbalance between the data size of the preictal periods and the interictal periods.

TABLE II
SEIZURE PREDICTION PERFORMANCE OF THE PROPOSED MODELS FOR MIT PHYSIONET SCALP EEG DATABASE

Patient	MLP				CNN				3D CNN			
	$S_n(\%)$	$\rho_w(\%)$	FPR(/h)	p	$S_n(\%)$	$\rho_w(\%)$	FPR(/h)	p	$S_n(\%)$	$\rho_w(\%)$	FPR(/h)	p
Pt 1	100.0±00.0	11.1±0.6	0.140±0.013	<0.001	100.0±00.0	7.7±0.7	0.035±0.017	<0.001	100.0±06.7	10.7±1.2	0.105±0.024	<0.001
Pt 2	66.7±00.0	31.3±2.4	0.729±0.056	**0.221**	100.0±00.0	7.9±2.5	0.133±0.053	<0.001	100.0±12.4	11.7±3.3	0.232±0.074	0.001
Pt 3	66.7±12.4	11.6±5.0	0.198±0.110	0.002	83.3±06.2	17.8±1.9	0.329±0.049	0.001	83.3±06.2	19.9±2.2	0.395±0.053	0.001
Pt 5	20.0±15.3	15.1±8.6	0.369±0.220	**0.546**	100.0±07.5	10.5±3.3	0.168±0.076	<0.001	100.0±07.5	8.3±1.1	0.101±0.030	<0.001
Pt 7	66.7±12.4	29.5±3.3	0.631±0.064	**0.199**	100.0±15.7	19.3±2.9	0.388±0.058	0.007	100.0±12.4	19.0±3.9	0.372±0.081	0.006
Pt 9	50.0±00.0	23.7±3.5	0.508±0.076	**0.229**	75.0±00.0	11.4±0.7	0.213±0.015	0.005	75.0±00.0	9.7±1.3	0.180±0.030	0.003
Pt 10	57.1±06.7	15.5±2.4	0.379±0.061	0.012	42.9±16.5	19.6±6.8	0.514±0.161	**0.132**	28.6±10.9	3.5±7.1	0.081±0.177	0.022
Pt 13	60.0±00.0	9.2±0.6	0.179±0.017	0.006	80.0±00.0	11.9±2.2	0.269±0.052	0.001	60.0±07.5	9.1±5.0	0.224±0.143	0.006
Pt 14	50.0±07.9	13.9±1.8	0.365±0.057	0.036	83.3±15.0	22.5±3.7	0.487±0.084	0.002	66.7±12.4	14.9±3.5	0.366±0.067	0.005
Pt 16	80.0±00.0	7.0±0.5	0.000±0.000	<0.001	80.0±00.0	8.2±1.2	0.000±0.051	<0.001	80.0±00.0	7.6±2.9	0.000±0.096	<0.001
Pt 17	66.7±00.0	1.3±2.0	0.000±0.047	<0.001	66.7±00.0	5.2±2.1	0.124±0.029	0.007	66.7±12.4	2.7±3.7	0.062±0.059	0.002
Pt 18	60.0±07.5	4.9±2.0	0.100±0.049	0.001	40.0±14.9	19.9±9.4	0.434±0.202	**0.248**	40.0±07.5	9.3±5.3	0.167±0.118	**0.067**
Pt 19	100.0±00.0	10.7±1.8	0.185±0.041	0.001	100.0±00.0	4.2±3.1	0.037±0.062	<0.001	100.0±00.0	12.0±2.4	0.222±0.051	0.002
Pt 20	83.3±07.9	5.5±4.2	0.000±0.069	<0.001	100.0±07.9	14.4±1.9	0.147±0.025	<0.001	100.0±08.3	11.7±1.7	0.098±0.037	<0.001
Pt 21	50.0±09.3	16.2±2.4	0.353±0.053	**0.118**	100.0±00.0	5.6±2.7	0.071±0.053	<0.001	100.0±00.0	11.5±1.2	0.212±0.026	<0.001
Pt 23	100.0±07.5	9.3±1.9	0.058±0.043	<0.001	100.0±00.0	7.8±0.9	0.000±0.027	<0.001	100.0±00.0	9.6±0.8	0.058±0.022	<0.001
Total	67.5±05.4	13.5±2.7	0.335±0.061	n.a	83.1±05.2	12.1±2.9	0.238±0.063	n.a	79.2±06.5	10.7±2.9	0.202±0.068	n.a

Significance levels of >0.05 are marked with bold.
S_n: sensitivity; ρ_w: proportion of time-in-warning; FPR: false prediction rate; p: one-sided p-value for superiority of the model compared to chance

To overcome this, we used random subsampling on the interictal data during the training phase. In this way, the rate of preictal and interictal data used during the training phase has become equal.

A multi-layer perceptron (MLP) model was created to compare with the performance of the proposed learning models This model consists of 3 hidden layers with 512 cells and an output layer. In order to prevent from overfitting during the training phase, 50% dropout was applied to entire hidden layers.

Our models are implemented in Python 2.7 with use of Lasagne 0.2 [44] with a Theano 1.0 [45] backend. The model is configured to run on NVIDIA GTX 1060 graphics card.

F. Post Processing

Successive alarms in a short period of time cause the false positive rate to increase. Therefore, a refractory period has been defined in which all subsequent alarms are ignored after the system is triggered. In this study, the refractory period was determined as 30 minutes. In addition, high-frequency variations were filtered by applying a 1-minute (30-window) causal moving average filter to the results of seizure prediction.

G. Obtaining Heatmaps of Critical Regions

The heat map of critical regions shows how significant the subregions in the image are in classification with CNN. Occlusion test is applied to obtain heat maps for each sample [46]. In this test, a square part of the image is occluded (i.e. set to 0) and then the net is tested for its propensity to predict the correct label. This process is repeated in all parts of the image to produce a heat map. At this stage, all color channels of image are occluded at the same time. The heatmaps are weighted with the prediction value of each sample and the average heatmap is generated for the preictal and ictal periods.

III. EXPERIMENTAL RESULTS

We have tested our learning models on scalp EEG of 16 patients with 77 seizure events in the CHB-MIT scalp EEG database. Seizure occurrence period (SOP) and seizure prediction horizon (SPH) definitions are utilized for seizure prediction [47]. SOP is defined as the period where the seizure is expected to arise; while SPH is defined as the duration between the alarm and the beginning of SOP. If there is no seizure in SOP after the alarm received in SPH, this is considered as a false prediction. Otherwise, the prediction is correct. In this study, SPH is considered as 1 minute and SOP as 30 minutes. In order to evaluate the performance of the proposed learning models, four parameters are used: sensitivity (S_n, the ratio of correctly predicted seizures to the total number of seizures), time-in-warning(ρ_w, proportion of time spent in warning), false-positive rate (FPR, the number of false alarms per hour), and p-value (significance of the improvement over chance).

Seizure prediction performances of the proposed models for the CHB-MIT scalp EEG database are shown in TABLE II. The MLP, CNN and 3D CNN models achieve a sensitivity of 67.33%, 83.12%, and 79.22% for the entire database, respectively. At these sensitivity values, the proportion of time spent in warning is 13.5%, 12.1%and 10.7% respectively for MLP, CNN, and 3D CNN models. Accordingly, the FPR values are obtained as 0.333/h, 0.238/h, and 0.202/h. The improvement-over-chance is statistically insignificant for five patients in the MLP model, two patients in the CNN model and one patient in the 3D CNN model, at a significance value of 0.05.

A. Effects of Epileptic Stage Lengths on Prediction Performance

There is no common proposition for epileptic stage lengths in the literature. The reason for this is that the duration of

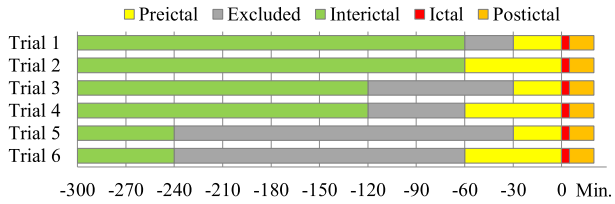


Fig. 3. Trials with various epileptic stage lengths. The trials were performed for the preictal period of 30 and 60 minutes and the interictal distance of 60, 120 and 240 minutes. The gray areas were left out of training.

epileptic stages varies according to the patient in fact. In the study, the preictal period was defined as 30 minutes and the distance between the interictal and ictal stages was 1 hour. The training performed with these periods has resulted in a lower seizure prediction performance for most patients. Therefore, all models were repeatedly trained for various epileptic stage lengths. As in Fig. 3, the trials were performed for the preictal period of 30 and 60 minutes and the interictal distance of 60, 120 and 240 minutes. Seizure prediction performances of the models trained with these periods are given in Fig. 4 with AUC values, which is a threshold-free metric. It has been observed that the increase in distance between the interictal and ictal stages leads to improvements in prediction performance of the models for most of the patients. For statistical confirmation of this observation, the significance of superiority of various epileptic stage lengths was obtained by Wilcoxon signed-rank test for AUC values. The significance values (p-value) matrix for CNN and 3D CNN models are given in Fig. 5. The superiority of the prediction performance obtained with the 240-minute interictal distance over the 60- and 120-minute interictal distance is statistically significant for 3D CNN model, at a significance level of 0.05.

B. Effects of Variable-Weighted Learning on Prediction Performance

It is known that the optimal preictal stage length varies according to the patient. For fixed preictal stage length, variable-weighted learning is suggested to minimize the effects of this problem. In the learning stage described in the previous section, all samples are represented with equal weight in the loss function. In the method proposed in this section, the positive (preictal) samples in the loss function are represented by a time-varying weight rather than a constant weight. For this purpose, a w_{pos} coefficient is added to the binary cross-entropy loss function given by (2) and the weighted representation of the loss function is provided as in (3). The value of this coefficient is determined depending on the location of the positive sample. It increases from the beginning to the end of the preictal stage, as given in (4). Here n_{pos} is the number of samples in the preictal stage. In this way, while the data at the beginning of the preictal stage is less effective in training, the data closer to the ictal stage will be more effective.

$$E_n = -\frac{1}{n}\sum_{k=1}^{n} w_{pos}.t_k.\log(y_k) + (1-t_k)\log(1-y_k) \quad (3)$$

$$w_{pos} = \frac{1}{2} + \frac{i_{pos}-1}{n_{pos}}, \quad i_{pos} = [1, 2, n_{pos}] \quad (4)$$

Seizure prediction performances of the 3D CNN models trained with 240-minute interictal distance and 60-minute preictal stage length are given in Table III comparatively for normal and variable-weighted learning. Although there is some increase in the overall sensitivity with the use of variable-weighted learning, there is also an increase in the proportion of time-in-warning. This suggests that variable weighted learning has no improving effect on overall seizure prediction performance.

C. Testing Against A Chance Predictor

In order to understand whether the performance of the proposed seizure prediction method is better than chance, the results were also compared with a chance predictor. Assuming that the interval between two successive alarms follows an exponential distribution as a Poisson process, the probability of rising of at least one alarm at random in an interval of Δt is approximately equal to $\lambda_w \Delta t$, independent of t, where λ_w is called the Poisson rate parameter. In this context, the sensitivity of the chance predictor, S_{nc}, is defined as in (5) [48]. The detection interval τ_{w0} corresponds to the seizure prediction horizon (SPH), while τ_w corresponds to the sum of SPH and the seizure occurrence period (SOP).

$$S_{nc} = 1 - exp\big(-\lambda_w \tau_w + (1 - e^{-\lambda_w \tau_{w0}})\big) \quad (5)$$

The difference between observed and chance sensitivity depend on ρ_w is a strong measure of predictability [48]. For a seizure prediction method with sensitivity S_n and proportion of time-in-warning ρ_w, the sensitivity improvement-over-chance metric is given in (6).

$$S_n - S_{nc} = S_n - 1 + exp\big(-\lambda_w \tau_w + (1 - e^{-\lambda_w \tau_{w0}})\big) \quad (6)$$

where

$$\lambda_w = -\frac{1}{\tau_w} ln(1 - \rho_w) \quad (7)$$

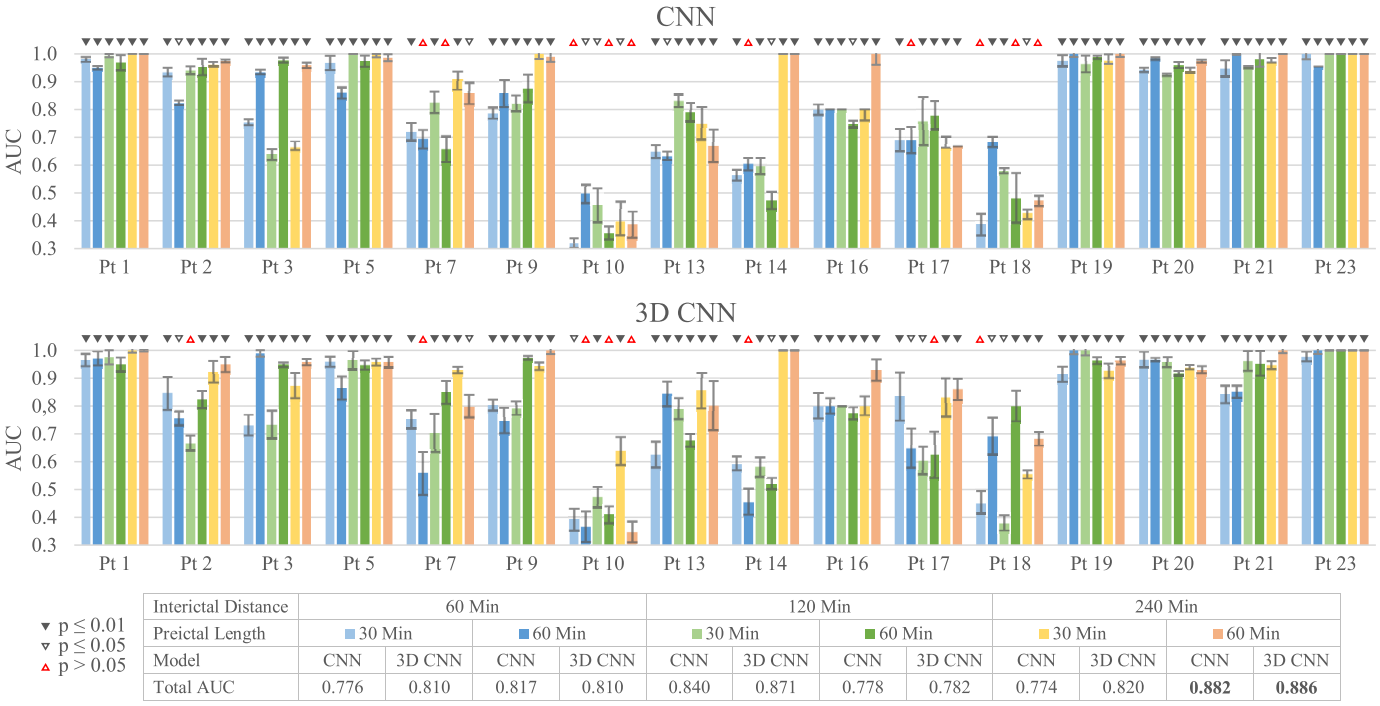
Assuming that the proposed prediction method correctly identifies n of N seizures for an individual patient, the significance of an improvement over chance is evaluated by the one-sided p-values, as given in (8) .

$$p = 1 - \sum_{i=0}^{n-1} \binom{N}{i} S_{nc}^{i} (1 - S_{nc})^{N-i}, \quad for \ \frac{n}{N} \geq S_{nc} \quad (8)$$

The significance of the sensitivity values is given in Table II, Fig. 4 and Table III with one-sided p values for each individual patient. Table III shows that the superiority of the proposed 3D CNN model to the chance predictor is statistically significant for all patients, except one patient, at a significance value of 0.05.

D. Heatmaps of Critical Regions

We also obtained heatmaps of critical regions in preictal and ictal periods for eight patients in this study where the primary goal was seizure prediction. The results obtained from



Interictal Distance	60 Min				120 Min				240 Min			
Preictal Length	30 Min		60 Min		30 Min		60 Min		30 Min		60 Min	
Model	CNN	3D CNN	CNN	3D CNN	CNN	3D CNN	CNN	3D CNN	CNN	3D CNN	CNN	3D CNN
Total AUC	0.776	0.810	0.817	0.810	0.840	0.871	0.778	0.782	0.774	0.820	**0.882**	**0.886**

Fig. 4. Seizure prediction performances of the proposed models with various epileptic stage lengths. Results are given with means and standard deviations of AUC values. Significance levels of ≤0.01, ≤0.05 and >0.05 are given with ▼, ▽, and △, respectively.

ID: Interictal Distance PL: Preictal Length

ID	PL	60 Min 30 Min	60 Min 60 Min	120 Min 30 Min	120 Min 60 Min	240 Min 30 Min	240 Min 60 Min
60 Min	30 Min	1	0.852	0.985	0.953	0.987	0.999
	60 Min	0.163	1	0.719	0.510	0.820	0.971
120 Min	30 Min	**0.018**	0.300	1	0.195	0.751	0.874
	60 Min	0.053	0.510	0.820	1	0.681	0.964
240 Min	30 Min	**0.016**	0.195	0.271	0.339	1	0.847
	60 Min	**0.000**	**0.034**	0.138	**0.042**	0.170	1

(a) CNN Model

ID: Interictal Distance PL: Preictal Length

ID	PL	60 Min 30 Min	60 Min 60 Min	120 Min 30 Min	120 Min 60 Min	240 Min 30 Min	240 Min 60 Min
60 Min	30 Min	1	0.489	0.479	0.908	0.999	0.999
	60 Min	0.533	1	0.572	0.943	0.985	0.988
120 Min	30 Min	0.541	0.452	1	0.757	0.998	0.995
	60 Min	0.101	0.064	0.259	1	0.954	0.982
240 Min	30 Min	**0.001**	**0.018**	**0.002**	0.052	1	0.798
	60 Min	**0.001**	**0.014**	**0.006**	**0.021**	0.212	1

(b) 3D CNN Model

Fig. 5. The significance values (p-value) matrix obtained by the one-tailed Wilcoxon signed-rank test reveals the superiority of different epileptic phase lengths (significance of the vertical axis to the horizontal axis) for (a) CNN model, (b) 3D CNN model. Significance levels of <0.05 are marked with bold.

a limited number of patients suggest that, the critical regions of the preictal and ictal periods are at different brain location, as in Fig. 6. The significance of this finding should be studied in detail on large patient groups.

IV. Discussion

Recently, several approaches have been proposed to predict the onset of epileptic seizures. The performance of our proposed method was presented in Table IV with the recent seizure prediction approaches using the same CHB-MIT scalp EEG database. It is difficult to determine which approach is better because each was tested with a limited set of data selected according to the defined time constraints and approach. Moreover, how generalizable is the proposed method is an important criterion that affects performance. Zhang and Parhi [13] proposed a patient-specific seizure prediction algorithm using spectral power features which ranked and selected in a patient-specific manner. With the use of patient-specific feature engineering, this approach achieved a high sensitivity of 98.68% and FPR of 0.046/h in the tests performed with 17 patients from the CHB-MIT scalp EEG dataset. However, this method, in which features are ranked and selected in a patient-specific manner, requires adequate knowledge and expertise for feature engineering in a new patient dataset.

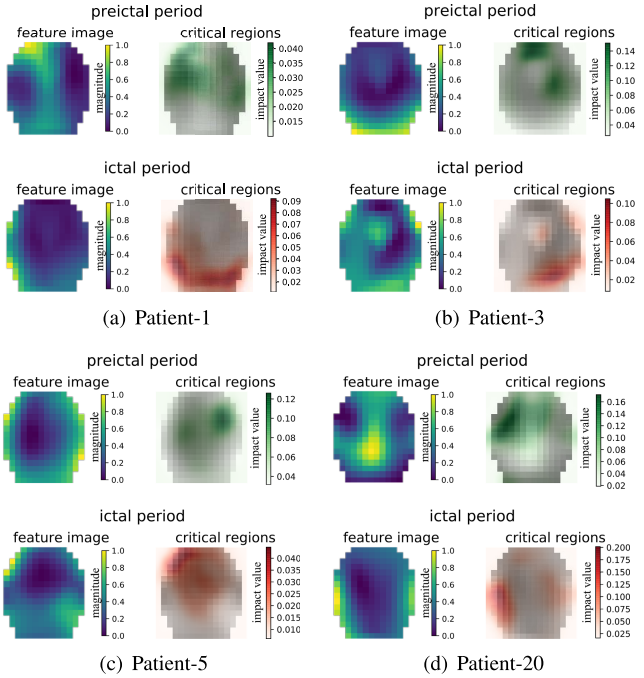
The number of seizures in the epileptic EEG data sets is limited. A total of 182 seizures occur during the 844 hour EEG recording of 23 patients in the CHB-MIT scalp EEG data set. Moreover, there are only 3 seizures that can be used for training in some patient datasets. The use of a limited number of seizures in prediction causes overfitting the training data. In order to reveal the seizure prediction performance of a proposed model under conditions similar to the real, leave-one-out cross-validation might be used for training. In this validation method, where all preictal data related to one of the seizures are excluded from the training and used only for testing, the performance is relatively low because no data with a high temporal dependence on the seizure data is used. Tsiouris *et al.* [50] proposed a seizure prediction model that employs the LSTM network. They have achieved a very high prediction sensitivity of 99.84% and low FPR of 0.02/h due to the fact that not using leave-one-out cross-validation.

In this paper, we proposed an approach for epileptic seizure prediction based on the use of multi-color image series created with features obtained from multichannel EEG signals as input to a 3D CNN. In the evaluation of the features generated from EEG data as images, we have been inspired by a

TABLE III
SEIZURE PREDICTION PERFORMANCE OF THE 3D CNN MODEL WITH NORMAL AND VARIABLE-WEIGHTED LEARNING

Patient	Normal				Variable-Weighted			
	$S_n(\%)$	$\rho_w(\%)$	FPR(/h)	p	$S_n(\%)$	$\rho_w(\%)$	FPR(/h)	p
Pt 1	100.0±00.0	11.3±2.0	0.000±0.026	<0.001	100.0±00.0	11.4±1.8	0.000±0.033	<0.001
Pt 2	100.0±00.0	14.9±4.3	0.154±0.058	0.003	100.0±00.0	17.6±4.4	0.192±0.053	0.005
Pt 3	100.0±00.0	17.1±1.3	0.110±0.017	<0.001	100.0±00.0	19.6±2.2	0.146±0.030	<0.001
Pt 5	100.0±07.5	10.2±2.6	0.139±0.062	<0.001	100.0±00.0	6.9±1.7	0.069±0.052	<0.001
Pt 7	66.7±15.7	10.7±3.6	0.136±0.056	0.031	66.7±16.7	7.3±4.7	0.078±0.071	0.015
Pt 9	100.0±00.0	4.9±1.2	0.000±0.023	<0.001	100.0±00.0	4.9±0.2	0.000±0.000	<0.001
Pt 10	42.9±09.8	18.1±2.2	0.456±0.061	**0.113**	57.1±13.5	18.7±6.7	0.414±0.149	0.025
Pt 13	80.0±00.0	5.6±3.8	0.265±0.113	<0.001	80.0±00.0	13.9±2.3	0.199±0.077	0.002
Pt 14	100.0±00.0	12.9±1.4	0.000±0.000	<0.001	100.0±00.0	12.9±1.3	0.000±0.000	<0.001
Pt 16	80.0±10.0	8.4±4.3	0.000±0.084	<0.001	100.0±14.9	11.2±4.3	0.000±0.084	<0.001
Pt 17	66.7±12.4	4.6±7.5	0.000±0.124	0.006	66.7±00.0	4.6±1.7	0.000±0.032	0.006
Pt 18	60.0±00.0	3.0±1.4	0.000±0.018	<0.001	60.0±10.0	20.8±4.7	0.303±0.080	**0.062**
Pt 19	100.0±00.0	6.4±1.9	0.037±0.017	<0.001	100.0±00.0	16.7±3.1	0.148±0.046	0.004
Pt 20	83.3±06.2	17.1±1.7	0.105±0.039	0.001	83.3±00.0	17.5±0.3	0.105±0.000	0.001
Pt 21	100.0±00.0	7.0±2.1	0.000±0.020	<0.001	100.0±00.0	5.7±2.1	0.000±0.029	<0.001
Pt 23	100.0±00.0	16.1±0.6	0.000±0.000	<0.001	100.0±00.0	16.2±0.6	0.000±0.000	<0.001
Total	85.7±03.9	10.5±2.6	0.096±0.045	n.a	88.3±03.4	12.9±2.6	0.116±0.046	n.a

Significance levels of >0.05 are marked with bold.
S_n: sensitivity; ρ_w: proportion of time-in-warning; FPR: false prediction rate; p: one-sided p-value for superiority of the model compared to chance



(a) Patient-1 (b) Patient-3

(c) Patient-5 (d) Patient-20

Fig. 6. Heatmaps of critical regions in preictal and ictal periods for (a) Patient-1, (b) Patient-3, (c) Patient-5, and (d) Patient-20. Feature images represent the average of 14 feature in the spatial domain throughout the entire preictal and ictal recordings in the patient concerned. The heatmap of critical regions shows how significant the regions in the image are in classification with CNN.

similar approach that Bashivan *et al.* [38] previously proposed for mental state classification from EEG data. In the tests performed with 16 patients from the CHB-MIT scalp EEG data set, we obtained a sensitivity rate of 85.7%, a false positive rate of 0.096/h, and a proportion of time spent in warning of 10.5% with the 3D CNN model. Table II shows that the proposed CNN and 3D CNN models achieve higher sensitivity with a lower proportion of time-in-warning and a lower FPR rate compared to the MLP model. During training of the proposed models, we used leave-one-out cross validation to ensure that the performance of the test data was similar to the real conditions. No patient-specific engineering was performed at any stage of this study.

In a similar approach, Truong *et al.* [11] proposed a CNN-based seizure prediction model using a short-term Fourier transform (STFT) in feature extraction to utilize both frequency and time aspects of EEG signals. In the tests performed with 13 patients from the same data set, they have achieved a sensitivity of 81.2% and FPR of 0.16/h in similar conditions with this work in terms of the epileptic stage lengths, cross-validation method, etc. In another seizure prediction approach, Khan *et al.* [10] proposed the use of wavelet transform of raw EEG signals as input to a CNN. In their study, a sensitivity of 87.8% and an FPR of 0.147/h were reported in the tests performed with 15 patients from the CHB-MIT scalp EEG data set. In order to directly compare the performance of these seizure prediction methods for each individual patient, the AUC values obtained for each patient are given comparatively in Fig. 7.

Our experiments with various timing constraints have shown us that epileptic stage lengths are an important factor affecting seizure performance. In both models, the best AUC values are obtained with a 60-minute preictal stage length and 240-minute interictal-ictal stage distance, as seen in Fig. 4. Although the increase in the distance between the interictal and ictal stage positively affects the seizure prediction performance, it causes a decrease in the total duration of the evaluated interictal phase. TABLE I shows the change of the total evaluated interictal data duration in hours according to the distance between the interictal and ictal stages.

We proposed variable-weighted learning to reduce the negative effects of the fixed-length preictal stage definition to prediction performance. The effect of variable-weighted learning on prediction performance is shown in Table III for 3D CNN model. Although the overall sensitivity was slightly increased,

TABLE IV
RESULTS OF RECENT EPILEPTIC SEIZURE PREDICTION APPROACHES ON CHB-MIT SCALP EEG DATASET

Authors	Dataset	Features	Classifier	No. of seizures	No. of subject	FPR(/h)	$S_n(\%)$	Interictal distance (minutes)	Preictal length (minutes)	Evaluated interictal hours
Zandi et al. [12]	MIT	Zero crossings, similarity/dissimilarity index	-	18	3	0.165	83.81	60	40	-
Zhang and Parhi [13]	MIT	Absolut/Relative spectral power	SVM	78	17	0.046	98.68	60	60	-
Myers et al. [49]	MIT	Phase/Amplitude locking value	-	31	10	0.17	77	-	60	31
Cho et al. [15]	MIT	Phase locking value	SVM	65	21	-	82.44	30	5	<10.83
Chu et al. [9]	MIT	Fourier Transform coefficients, PSD	-	125	13	0.392	83.33	-	86	<434.6
Alotaiby et al. [14]	MIT	Common spatial pattern statistics	LDA	170	24	0.47 0.4 0.39	81 87 89	-	60 90 120	<982.9
Truong et al. [11]	MIT	STFT spectral images	CNN	64	13	0.16	81.2	240	30	209
Khan et al. [10]	MIT	Wavelet transform coefficient	CNN	18	15	0.147	87.8	-	10	<70,5
Tsiouris et al. [50]	MIT	Statistical moments, zero crossings, Wavelet transform coefficients, PSD	LSTM	185	24	0.11 0.06 0.03 0.02	99.28 99.38 99.63 99.84	-	15 30 60 120	<979.9
This work	MIT	Spectral power, Statistical moments, Hjorth parameters	3D CNN	77	16	0,292 0,186 0,096	86,84 87,01 85,71	60 120 240	60 60 60	466,1 419,4 353,5

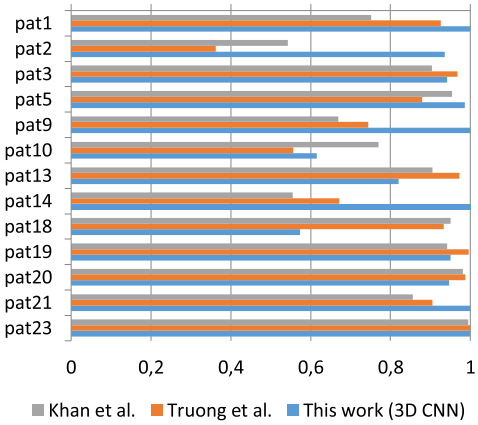


Fig. 7. Comparison of AUC values for individual patients in the MIT scalp EEG dataset among Khan et. al., Truong et. al., and this work. AUC, area under the receiver operating characteristic curve.

the increase in the proportion of time-in-warning indicates that variable-weighted learning did not positively affect total seizure prediction performance.

In order to understand whether the performance of the proposed methods is better than chance, the results were also compared with a Poisson based chance predictor. The results show that the superiority of the proposed 3D CNN model to chance is statistically significant for all patients, except Patient-10, at a significance value of 0.05.

V. CONCLUSION

This paper presents a CNN-based patient-specific seizure prediction method in which features obtained from multichannel EEG signals are evaluated as a multi-color image series. Spectral band power, statistical moment and Hjorth parameters were used to reveal the frequency and time domain features of the EEG signals. The proposed 3D CNN model achieved a sensitivity of 85.7%, a false prediction rate of 0.096/h, and a proportion of time-in-warning of 10.5%, in the tests performed with 16 patients from the CHB-MIT scalp EEG data set. The results show that the seizure prediction approach proposed in this study performs close to the state-of-the-art approaches under similar conditions. Our proposed method is generalizable due to the fact that no patient-specific engineering is performed at any stage. However, the method needs to be comprehensively tested with more subjects under different clinical conditions from different age groups to determine the overall performance because of the CHB-MIT scalp EEG data set is mostly composed of pediatric patients.

In order to overcome the negative effects of the fixed-length preictal period definition on prediction performance, variable-weighted learning has been proposed in this work, however, no improving effect has been obtained.

Our experiments with various durations show that accurate determination of epileptic stage lengths according to each individual patient is an important criterion affecting seizure prediction success. Furthermore, the results of these trials reveal that the increase in the distance between the interictal and ictal phases positively affects seizure prediction performance. This suggests that some preictal activities may occur even a few hours before seizure onset.

REFERENCES

[1] A. Yadollahpour and M. Jalilifar, "Seizure prediction methods: A review of the current predicting techniques," *Biomed. Pharmacol. J.*, vol. 7, no. 1, pp. 153–162, 2015. [Online]. Available: http://www.biomedpharmajournal.org/absdoic.php?snoid=466

[2] D. R. Freestone, P. J. Karoly, and M. J. Cook, "A forward-looking review of seizure prediction," *Current Opinion Neurol.*, vol. 30, no. 2, pp. 167–173, 2017.

[3] S. B. Dumanis, J. A. French, C. Bernard, G. A. Worrell, and B. E. Fureman, "Seizure forecasting from idea to reality. Outcomes of the my seizure gauge Epilepsy innovation institute workshop," *eNeuro*, vol. 4, no. 6, pp. 1–5, Nov./Dec. 2017.

[4] M. O. Baud and V. R. Rao, "Gauging seizure risk," *Neurology*, vol. 91, no. 21, pp. 967–973, 2018.

[5] L. Kuhlmann, K. Lehnertz, M. P. Richardson, B. Schelter, and H. P. Zaveri, "Seizure prediction—Ready for a new era," *Nature Rev. Neurol.*, vol. 14, pp. 618–630, Aug. 2018.

[6] W. C. Stacey, "Seizure prediction is possible–now let's make it practical," *EBioMedicine*, vol. 27, pp. 3–4, Jan. 2018.

[7] P. R. Carney, S. Myers, and J. D. Geyer, "Seizure prediction: Methods," *Epilepsy Behav.*, vol. 22, pp. S94–S101, Dec. 2011. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1525505011005087

[8] M. Bandarabadi, C. A. Teixeira, J. Rasekhi, and A. Dourado, "Epileptic seizure prediction using relative spectral power features," *Clin. Neurophysiol.*, vol. 126, no. 2, pp. 237–248, 2015. doi: 10.1016/j.clinph.2014.05.022.

[9] H. Chu, C. K. Chung, W. Jeong, and K.-H. Cho, "Predicting epileptic seizures from scalp EEG based on attractor state analysis," *Comput. Methods Programs Biomed.*, vol. 143, pp. 75–87, May 2017.

[10] H. Khan, L. Marcuse, M. Fields, K. Swann, and B. Yener, "Focal onset seizure prediction using convolutional networks," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 9, pp. 2109–2118, Sep. 2017.

[11] N. D. Truong *et al.*, "Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram," *Neural Netw.*, vol. 105, pp. 104–111, Sep. 2018. doi: 10.1016/j.neunet.2018.04.018.

[12] A. S. Zandi, R. Tafreshi, M. Javidan, and G. A. Dumont, "Predicting epileptic seizures in scalp EEG based on a variational Bayesian Gaussian mixture model of zero-crossing intervals," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 5, pp. 1401–1413, May 2013.

[13] Z. Zhang and K. K. Parhi, "Low-complexity seizure prediction from iEEG/sEEG using spectral power and ratios of spectral power," *IEEE Trans. Biomed. Circuits Syst.*, vol. 10, no. 3, pp. 693–706, Jun. 2016.

[14] T. N. Alotaiby, S. A. Alshebeili, F. M. Alotaibi, and S. R. Alrshoud, "Epileptic seizure prediction using CSP and LDA for scalp EEG signals," *Comput. Intell. Neurosci.*, vol. 2017, 2017, Art. no. 1240323.

[15] D. Cho, B. Min, J. Kim, and B. Lee, "EEG-based prediction of epileptic seizures using phase synchronization elicited from noise-assisted multivariate empirical mode decomposition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 8, pp. 1309–1318, Aug. 2017.

[16] M. J. Cook *et al.*, "Prediction of seizure likelihood with a long-term, implanted seizure advisory system in patients with drug-resistant epilepsy: A first-in-man study," *Lancet Neurol.*, vol. 12, no. 6, pp. 563–571, 2013.

[17] I. Kiral-Kornek *et al.*, "Epileptic seizure prediction using big data and deep learning: Toward a mobile system," *EBioMedicine*, vol. 27, pp. 103–111, Dec. 2017.

[18] F. Mormann *et al.*, "On the predictability of epileptic seizures," *Clin. Neurophysiol.*, vol. 116, no. 3, pp. 569–587, Mar. 2005.

[19] Y. Park, L. Luo, K. K. Parhi, and T. Netoff, "Seizure prediction with spectral power of EEG using cost-sensitive support vector machines," *Epilepsia*, vol. 52, no. 10, pp. 1761–1770, 2011.

[20] L. Chisci *et al.*, "Real-time epileptic seizure prediction using AR models and support vector machines," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 5, pp. 1124–1132, May 2010.

[21] B. Litt *et al.*, "Epileptic seizures may begin hours in advance of clinical onset: A report of five patients," *Neuron*, vol. 30, no. 1, pp. 51–64, Apr. 2001.

[22] F. Mormann, R. G. Andrzejak, C. E. Elger, and K. Lehnertz, "Seizure prediction: The long and winding road," *Brain*, vol. 130, no. 2, pp. 314–333, 2006.

[23] L. D. Iasemidis *et al.*, "Adaptive epileptic seizure prediction system," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 5, pp. 616–627, May 2003.

[24] K. Lehnertz and C. E. Elger, "Can epileptic seizures be predicted? Evidence from nonlinear time series analysis of brain electrical activity," *Phys. Rev. Lett.*, vol. 80, no. 22, pp. 5019–5022, 1998.

[25] M. Le Van Quyen *et al.*, "Anticipation of epileptic seizures from standard EEG recordings," *Lancet*, vol. 357, no. 9251, pp. 183–188, Jan. 2001.

[26] L. Kuhlmann *et al.*, "Patient-specific bivariate-synchrony-based seizure prediction for short prediction horizons," *Epilepsy Res.*, vol. 91, nos. 2–3, pp. 214–231, 2010.

[27] S. Taran, V. Bajaj, and S. Siuly, "An optimum allocation sampling based feature extraction scheme for distinguishing seizure and seizure-free EEG signals," *Health Inf. Sci. Syst.*, vol. 5, no. 1, p. 7, Dec. 2017.

[28] A. R. Hassan, S. Siuly, and Y. Zhang, "Epileptic seizure detection in EEG signals using tunable-Q factor wavelet transform and bootstrap aggregating," *Comput. Methods Programs Biomed.*, vol. 137, pp. 247–259, Dec. 2016.

[29] H. R. Al Ghayab, Y. Li, S. Siuly, and S. Abdulla, "A feature extraction technique based on tunable Q-factor wavelet transform for brain signal classification," *J. Neurosci. Methods*, vol. 312, pp. 43–52, Jan. 2019.

[30] K. Fei, W. Wang, Q. Yang, and S. Tang, "Chaos feature study in fractional Fourier domain for preictal prediction of epileptic seizure," *Neurocomputing*, vol. 249, pp. 290–298, Aug. 2017.

[31] H. R. Al Ghayab, Y. Li, S. Siuly, and S. Abdulla, "Epileptic seizures detection in EEGs blending frequency domain with information gain technique," *Soft Comput.*, vol. 23, no. 1, pp. 227–239, 2019.

[32] H.-T. Shiao *et al.*, "SVM-based system for prediction of epileptic seizures from iEEG signal," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 5, pp. 1011–1022, May 2017.

[33] M. Behnam and H. Pourghassem, "Real-time seizure prediction using RLS filtering and interpolated histogram feature based on hybrid optimization algorithm of Bayesian classifier and Hunting search," *Comput. Methods Programs Biomed.*, vol. 132, pp. 115–136, Aug. 2016. doi: 10.1016/j.cmpb.2016.04.014.

[34] A. R. Özcan and S. Ertürk, "Epileptic seizure prediction with recurrent convolutional neural networks," in *Proc. 25th Signal Process. Commun. Appl. Conf. (SIU)*, May 2017, pp. 1–4.

[35] A. H. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," Ph.D. dissertation, Harvard-MIT Health Sci. Technol., Massachusetts Inst. Technol., Cambridge, MA, USA, 2009.

[36] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. 215–220, 2000.

[37] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalogr. Clin. Neurophysiol.*, vol. 29, no. 3, pp. 306–310, 1970.

[38] P. Bashivan, I. Rish, M. Yeasin, and N. Codella, "Learning representations from EEG with deep recurrent-convolutional neural networks," pp. 1–15, 2015, *arXiv:1511.06448*. [Online]. Available: http://arxiv.org/abs/1511.06448

[39] J. P. Snyder, *Map Projections—A Working Manual*. Washington, DC, USA: U.S. Government Publishing Office, 1987.

[40] P. Alfeld, "A trivariate clough—Tocher scheme for tetrahedral data," *Comput. Aided Geometric Des.*, vol. 1, no. 2, pp. 169–181, 1984.

[41] P. W. Mirowski, Y. LeCun, D. Madhavan, and R. Kuzniecky, "Comparing SVM and convolutional networks for epileptic seizure prediction from intracranial EEG," in *Proc. IEEE Workshop Mach. Learn. Signal Process. (MLSP)*, 2008, pp. 244–249.

[42] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," pp. 1–10, 2014, *arXiv:1409.1556*. [Online]. Available: https://arxiv.org/abs/1409.1556

[43] X. Cao, "A practical theory for designing very deep convolutional neural networks," Tech. Rep., 2015.

[44] S. Dieleman *et al.*, "Lasagne: First release," Tech. Rep., 2015.

[45] The Theano Development Team, "Theano: A Python framework for fast computation of mathematical expressions," May 2016, *arXiv:1605.02688*. [Online]. Available:https://arxiv.org/abs/1605.02688

[46] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Computer Vision—ECCV* (Lecture Notes in Computer Science), vol. 8689. Zürich, Switzerland: Springer, 2014, pp. 818–833.

[47] T. Maiwald, M. Winterhalder, R. Aschenbrenner-Scheibe, H. U. Voss, A. Schulze-Bonhage, and J. Timmer, "Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic," *Phys. D, Nonlinear Phenomena*, vol. 194, nos. 3–4, pp. 357–368, 2004.

[48] D. E. Snyder, J. Echauz, D. B Grimes, and B. Litt, "The statistics of a practical seizure warning system," *J. Neural Eng.*, vol. 5, no. 4, p. 392, 2008.

[49] M. H. Myers, A. Padmanabha, G. Hossain, A. L. de Jongh Curry, and C. D. Blaha, "Seizure prediction and detection via phase and amplitude lock values," *Frontiers Hum. Neurosci.*, vol. 10, p. 80, Mar. 2016. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4781861

[50] K. M. Tsiouris, V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis, "A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals," *Comput. Biol. Med.*, vol. 99, pp. 24–37, Aug. 2018.