Epileptic EEG Classification by Using Time-Frequency Images for Deep Learning

Article *in* International Journal of Neural Systems · August 2021

DOI: 10.1142/S012906572150026X

3 authors:

Mehmet Akif Özdemir
Harvard Medical School
58 PUBLICATIONS **762** CITATIONS

SEE PROFILE

Özlem Karabiber Cura
Izmir Katip Celebi University
45 PUBLICATIONS **290** CITATIONS

SEE PROFILE

Aydin Akan
Izmir University of Economics
397 PUBLICATIONS **3,322** CITATIONS

SEE PROFILE

International Journal of Neural Systems, Vol. 31, No. 8 (2021) 2150026 (16 pages)

DOI: 10.1142/S012906572150026X

World Scientific
www.worldscientific.com

Epileptic EEG Classification by Using Time-Frequency Images for Deep Learning

Mehmet Akif Ozdemir* and Ozlem Karabiber Cura†
Department of Biomedical Engineering
Izmir Katip Celebi University
Cigli 35620, Izmir, Turkey
**makif.ozdemir@ikcu.edu.tr*
†*ozlem.karabiber@ikcu.edu.tr*

Aydin Akan‡
Department of Electrical and Electronics Eng.
Izmir University of Economics
Balcova 35330, Izmir, Turkey
akan.aydin@ieu.edu.tr

Received 2 April 2021
Accepted 3 April 2021
Published Online 26 May 2021

Epilepsy is one of the most common brain disorders worldwide. The most frequently used clinical tool to detect epileptic events and monitor epilepsy patients is the EEG recordings. There have been proposed many computer-aided diagnosis systems using EEG signals for the detection and prediction of seizures. In this study, a novel method based on Fourier-based Synchrosqueezing Transform (SST), which is a high-resolution time-frequency (TF) representation, and Convolutional Neural Network (CNN) is proposed to detect and predict seizure segments. SST is based on the reassignment of signal components in the TF plane which provides highly localized TF energy distributions. Epileptic seizures cause sudden energy discharges which are well represented in the TF plane by using the SST method. The proposed SST-based CNN method is evaluated using the IKCU dataset we collected, and the publicly available CHB-MIT dataset. Experimental results demonstrate that the proposed approach yields high average segment-based seizure detection precision and accuracy rates for both datasets (IKCU: 98.99% PRE and 99.06% ACC; CHB-MIT: 99.81% PRE and 99.63% ACC). Additionally, SST-based CNN approach provides significantly higher segment-based seizure prediction performance with 98.54% PRE and 97.92% ACC than similar approaches presented in the literature using the CHB-MIT dataset.

Keywords: Convolutional Neural Network (CNN); Deep Learning (DL); seizure detection; seizure prediction; segment-based; Synchrosqueezed Transform (SST); time-frequency images.

1. Introduction

Epilepsy is one of the most widespread brain disorders, affecting approximately 50 million people worldwide, characterized by abnormal electrical activities of the brain nerve cells resulting in recurrent seizures, unusual behavior, and maybe the loss of consciousness.[1] In order to decrease the harm of unexpected seizures and to start treatment as soon as possible, a timely and correct diagnosis of epilepsy is very compulsory for patients. In clinical practice, the diagnosis of epilepsy is made using detailed patient history, neurological examination,

‡Corresponding author.

and various clinical tools such as neuroimaging and EEG recording. However, achieving the correct diagnosis by visual examination of long-term EEG records or neurological images is directly related to the experience of the expert neurologist, and it is a very time-consuming and exhausting process.[2] Therefore, effective computer-aided diagnosis systems have been developed by researchers for many years to accurately and quickly detect and predict seizure activity.[1–3] It is possible to categorize the studies in this area into two major groups: (i) machine learning (ML)-based and (ii) deep learning (DL)-based approaches. Many ML-based methods have been proposed for automatic seizure detection, prediction, and classification, generally based on advanced signal processing and feature extraction algorithms. Time-domain and spectral-domain features have been utilized for automatic classification of seizure activities. In addition, various types of entropy such as fuzzy entropy,[4] approximate entropy, sample entropy, and phase entropy[5] have been computed from the EEG signals to identify different epileptic EEG segments. Hence, encouraging seizure detection performances have been achieved using the combination of various features calculated from either signal domain[6] or transform domains such as EMD and its derivatives,[7–9] Fourier transform (FT) and derivatives,[10,11] and Wavelet transform (WT) and derivatives.[1,3] Besides, employing time-domain features, or frequency-domain features, or their combination yields limited performance, due to nonstationary properties of EEG signals. Therefore, many time-frequency representation (TFR) methods have been utilized to improve seizure detection, prediction, and classification performances.[12–14] The most frequently used TFR methods are Wigner–Ville Distribution (WVD),[12] smoothed WVD,[15] Choi–Williams Distribution (CWD), Modified-B Distribution (MBD),[13] Short-Time Fourier Transform (STFT),[16] and WT.[17,18] A recently developed, high-resolution TFR method called SST has also been applied into the epileptic EEG analysis.[10,11,19] Image-based features,[11,17,19] or TF-domain features which are the extended versions of time-, or frequency-domain features,[12,13,19] are extracted from the calculated TFRs. Then various classifiers such as k-nearest neighbor,[6] support vector machines,[11] artificial neural networks,[9] Nave Bayes,[19] random forest classifiers,[7] and extreme learning machines[20] are employed to identify seizure and seizure-free EEG segments by extracted features.

As for the second category, successful DL-based methods have been proposed for seizure detection and prediction purposes. ML-based methods are conducted by extracting various features, performing statistical analysis, and using various classifiers to achieve the best classification performance; besides, DL-based methods do not require manual feature extraction and selection processes.[2] In DL-based approaches, multi-channel EEG signals, their combinations,[2,21–25] or image-based representations of the EEG signals obtained using TFR methods[26–29] are utilized as input to train a deep network. For example, in a study,[2] a 13-layer deep CNN-based epileptic seizure classification approach is introduced to distinguish three EEG groups, normal, pre-seizure, and seizure with a classification accuracy of 95.00%. Ansari *et al.* proposed a 23-layer CNN-based method to classify seizure and nonseizure neonatal EEG signal segments. They compared the performance of the proposed CNN-based approach with feature-based, and heuristic methods that did not contain a complex processing stage.[21] The deep CNN-based epileptiform discharges (ED) detection approach was introduced in another study. EEG data without ED of 25 epilepsy patients, and 25 age-matched control EEG were used to train a 13-layer CNN with promising detection accuracies.[22] In another study, a seizure control method was proposed by utilizing the combination CNN model where prior to the convolution operation, the combinations of channels are obtained using multi-channel EEG recordings.[23] Li *et al.* introduced a CNN-based automatic seizure detection approach that contains the Nested Long Short-Term Memory (NLSTM) model. Without requiring any advanced signal processing steps, high seizure detection performances were obtained on three different datasets.[24] In another study, a novel seizure onset detection approach was introduced using the Stockwell transform and 15-layer deep CNN, resulting higher than 95.45% sensitivities.[26]

In this study, we propose an image-based utilization of TFRs obtained by the high-resolution SST of EEG signals. SST approach provides a close to ideal TF energy distribution by assigning the TF components of the signal into the instantaneous frequency (IF) trajectory and resulting in very high

TF localization of the signal components. Epileptic seizures cause abrupt changes on the statistics of EEG signals which are nonstationary in nature. Hence, the aim of this study is to apply a high-resolution TF analysis method such as SST, into the representation of epileptic EEG signals to be able to capture the fast spectral variations of EEG signals during seizures. We train a CNN architecture with these TF images to achieve high segment-based seizure detection and prediction performances. The performance evaluation of both seizure detection and prediction is based on a segment-based method which consists of 1 and 5 s long EEG segments instead of the whole seizure event. The proposed approach is tested on two datasets: IKCU dataset we collected and the CHB-MIT dataset to detect and predict epileptic seizures with outstanding validation accuracies.

2. Classification of Epileptic EEG based on Time-Frequency Image

In this paper, a novel SST-based DL approach is introduced to discriminate seizure and other (pre-seizure or inter seizure) EEG segments. The framework of the proposed study is shown in Fig. 1. The proposed approach includes achieving joint TF representations of seizure and pre-seizure (or inter-seizure) EEG segments marked by the experts by utilizing the SST method. We use the resulting highly localized TF images to train a CNN. The applications of CNN have gained popularity in recent years due to the high success of seizure detection, prediction, and classification, without the need for complex calculations, additional feature extraction and classification processes.[2] In the following, the SST method utilized in our study to obtain TF images of EEG segments is briefly introduced.

2.1. *Synchrosqueezing transform*

Synchrosqueezing Transform is a special member of the *reallocation* methods family, and it was initially introduced to analyze auditory signals. The main purpose of this method is achieving a highly localized TFR by reassigning the value of TF coefficients (ω, t), obtained by another TF analysis method, to a distinct point $(\omega_0(\omega, t), t)$, where $\omega_0(\omega, t)$ is the local IF. In order to achieve high-resolution TFRs of nonstationary signals, SST algorithms based on both CWT and STFT may be utilized.[11,30,31] In our proposed approach, Fourier (STFT)-based SST is used to generate TF images for CNN classification. The STFT-based SST algorithm starts by calculating the STFT "$X(\omega, t)$" of the signal "$x(t)$" to be analyzed.

$$X(\omega, t) = \int_{-\infty}^{\infty} x(\tau) w(\tau - t) \mathrm{e}^{-j\omega\tau} d\tau, \qquad (1)$$

where $w(t)$ indicates the window function. Then, the IF of the signal "$\omega_0(\omega, t)$" is estimated by calculating

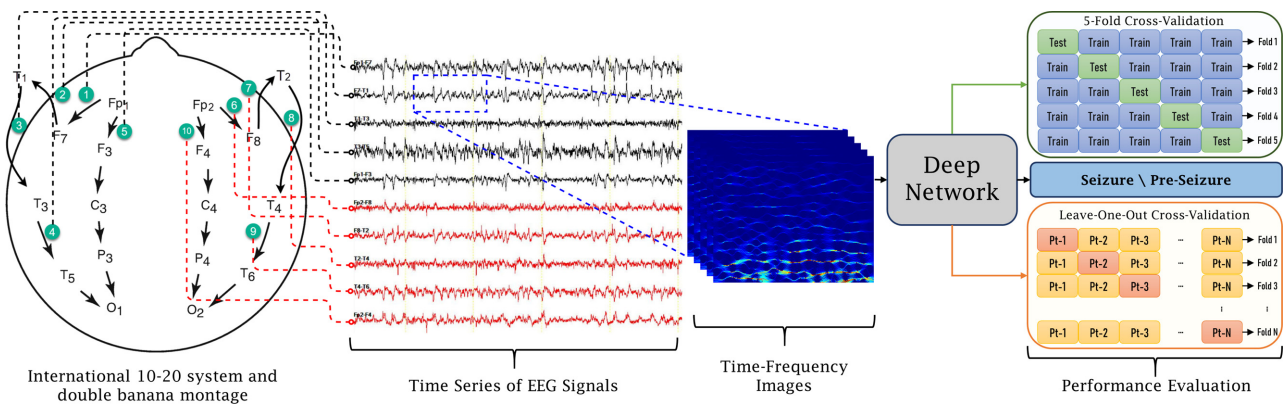


Fig. 1. Proposed study framework; EEG signal acquisition with the international 10-20 system and double banana montage, EEG time-series segmentation, generating TF images, and training and testing the deep network. Both patient-correlated 5-fold cross-validation and patient-independent leave-one-out cross-validation methods have been used for the robustness evaluation of the performance of the deep network ($N = 16$ for IKCU dataset and $N = 23$ for CHB-MIT dataset).

the derivative of "$X(\omega, t)$" with respect to time.

$$\omega_0(\omega, t) = \frac{1}{2\pi}\partial_t \arg X(\omega, t) = \mathrm{Re}\left(\frac{1}{2i\pi}\frac{\partial_t X(\omega, t)}{X(\omega, t)}\right). \quad (2)$$

SST "$T(\eta, t)$" can be formulated using vertical reconstruction formula and squeezing operator $\int_{-\infty}^{\infty} \delta(\eta - \omega_0(\omega, t))d\omega$, and the STFT coefficients that have the same frequency details are reassigned to in which they should appear.[11,31]

$$T(\eta, t) = \frac{1}{g(0)} \int_R X(\omega, t)\delta(\eta - \omega_0(\omega, t))d\omega. \quad (3)$$

To reveal the advantages of the proposed SST approach, we also performed the traditional TF analysis method, STFT and used in our CNN-based seizure classification. Example TF representations calculated using both the proposed SST approach and the STFT of pre-seizure and seizure EEG segments are given in Fig. 2. In our study, the window function that will provide the best possible resolution for the STFT was chosen in order to compare it with SST. Nevertheless, improved TF resolutions are achieved for both pre-seizure and seizure EEG segments using the SST approach than that of STFT, as observed in Figs. 2(c)–2(f).

2.2. *Proposed deep CNN architecture*

In the proposed study, a novel network architecture is adapted from the 50-layer Residual Neural Network (ResNet-50) model, which is an advanced CNN architecture. ResNet-based approaches have been developed to prevent degradation and vanishing gradients problems as the network's depth increases in CNN architectures.[32] ResNet architecture adds shortcuts between layers to overcome this problem. In our proposed method, TFR images are used as input to train the network. Using TFR images, the energy density of the signals may be expressed as color pixels. In order to achieve high classification performance, all information in TF pixels must be processed with minimum loss. The information in TF energy density of nonstationary signals is critical when analyzing these signals.[33] In recent years, TFR images have been used frequently for training DL-based architectures.[34,35] Hence, in this study, a new CNN architecture is designed inspired by ResNet-50 architecture to classify the calculated TF images. The original ResNet-50 architecture has widely been used in

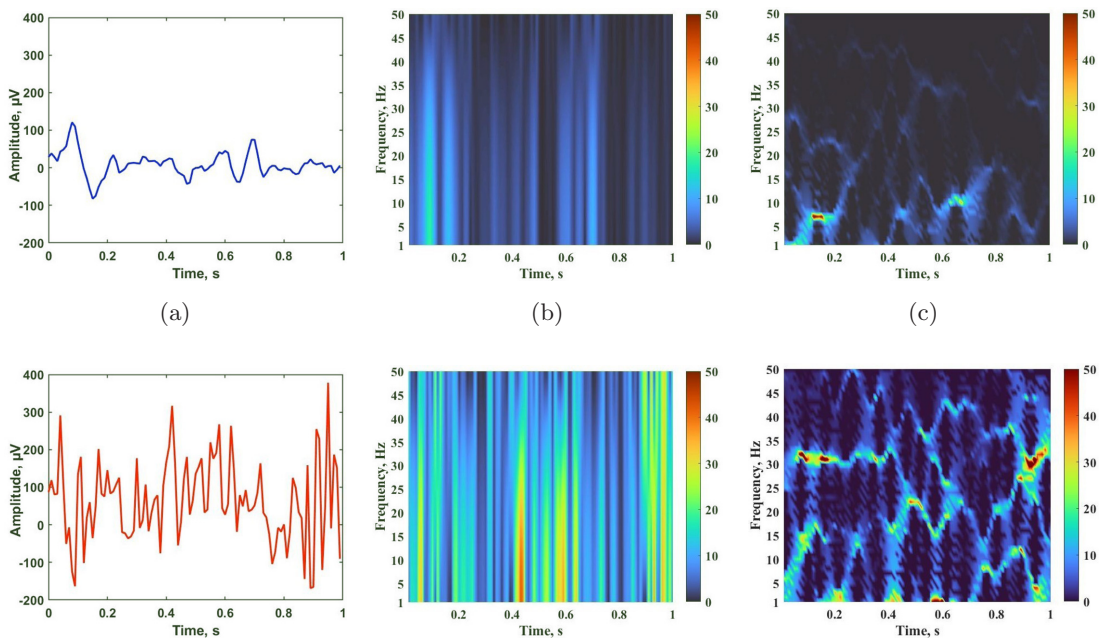


Fig. 2. The 1 s-duration pre-seizure of (a) EEG segment example, (b) STFT magnitude, and (c) SST magnitude; and 1 s-duration seizure of (d) EEG segment example, (e) STFT magnitude, and (f) SST magnitude.

many studies with successful results.[35] Nevertheless, it may be further optimized for binary classification problems with less training set. Hence, in this study, we use an optimized version of the ResNet-50. The proposed architecture includes fewer filters than the ResNet-50 and aims to increase the accuracy of the binary classification problem compared to the ResNet-50 while reducing the training cost.

A graphical representation of the ResNet-50-based proposed architecture and the output dimensions of the network layers are shown in Fig. 3. First, the TF-images recorded at 500 DPI resolution are resized to $128 \times 128 \times 3$ dimensions to reduce the training cost and are given as input to the deep network. Then, zero-padding is performed to balance the kernel sizes. Proposed architecture was implemented in five stages. In the first stage, one convolutional layer (including; convolution (Conv) (7×7), strides of 2, batch normalization (BN), and rectified linear unit (ReLU) function as the activation function) and one pooling layer (maximum) (3×3) strides of 2 have been used, respectively. The second stage consists of two blocks called convolutional block and identity block, both consisting of three repetitive convolutional layers and a step called shortcut that adds the incoming unaltered input array to the output of the convolutional layers. In the convolutional block, unlike in the identity block, the input array participates in the adding step after passing through Conv and BN layers on the shortcut phase. In the convolutional and identity blocks, the convolutional layers have 1×1, 3×3, and 1×1 kernel sizes, respectively. In the shortcut phase of the convolutional block, the convolutional layer has a 1×1 kernel size. Shortcuts have been used for preventing the deep network from overfitting, and also used to reduce and optimize computational complexity. After the second stage, the identity block is used by repeating 2, 3, 4, and 5 times, respectively. The proposed architecture includes 50-layer similar to the ResNet-50 architecture. The proposed architecture was not added more layers in order not to increase the training cost. It is modified for the binary classification problem by reducing the number of filters in the convolutional and identity blocks. The filter numbers in repetitive convolutional and identity blocks are as follows: while Stage-1 is preserved exactly as in ResNet-50, in Stage-2, convolutional block and all identity blocks have 16, 32, and 64 filters, respectively. The number of filters is increased in sub-stages so as to extract deeper features. Stage-3 has 32, 64, 128, and 128 filters. In all sub-stages of Stage-4, similar to the ResNet-50 architecture the number of filters kept fixed but reduced to 128. Similarly, the number of filters is reduced, chosen as 256, and fixed in all sub-stages of Stage-5. Deep residual features are extracted from the last stage. An average pooling (2×2) process was performed before the fully connected layer with 8192 neurons to reduce the large deep residual features array. Finally, the SoftMax function[36] has been used as a binary classifier to classify deep residual features. The deep network has over 23.5 million trainable parameters.

During the training phase of the network, Adam Optimizer was used, due to the effective choice of hyperparameters. Furthermore, different batch sizes have been tested in the training phase to obtain the best training result, and the batch size has been

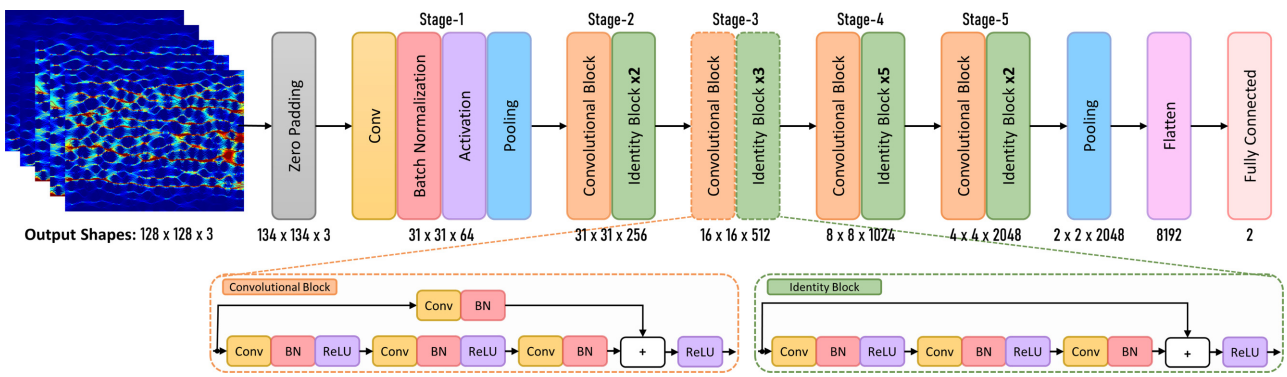


Fig. 3. Graphical representation of ResNet-50-based proposed CNN architecture (Conv: Convolution, BN: Batch Normalization, ReLU:Rectified Linear Unit).

optimized to 64 as parameter tuning. Epochs have been selected as 50 in order to specify standards in all training phases, not to increase the training cost, and to observe the robustness of the model. The learning rate was chosen as 0.1, 0.01, 0.001, and 0.0001, to observe its effect.

2.3. *Clinical EEG datasets*

The proposed CNN-based seizure detection and prediction approach combining with the SST is conducted with two different epileptic EEG datasets. First, epileptic EEG dataset **(IKCU dataset)** recorded at Izmir Katip Celebi University School of Medicine, Department of Neurology are examined for the proposed study after obtaining ethical approval (ethical approval dated 08.08.2019 and numbered 2969). This epileptic EEG dataset consists of 18-channel EEG recordings of 5 females and 11 males with a total of 16 epilepsy patients with an average age of 37.3 ∓ 7 (507 s and 892 s for pre-seizure and seizure, respectively). EEG recordings are obtained using the standard International 10-20 electrode system in a total of 18 channels with a sampling frequency of 100 Hz. Since it is known that epileptic seizures are predominant in the left temporal and frontal lobes for these patients, temporal and frontal lobes weighed 10-EEG channels (Fp1-F7, F7-T1, T1-T3, T3-T5, Fp1-F3, Fp2-F8, F8-T2, T2-T4, T4-T6, Fp2-F4) are used in our study. 1 s and 5 s long segment durations are tested in the proposed *segment-based seizure detection* study to calculate the TFRs using both SST and STFT methods.

Another dataset used in the proposed study is the publicly available 23/18 channel EEG recordings collected at Childrens Hospital Boston **(CHB-MIT dataset)**.[37] This dataset includes EEG recordings of 22 pediatric patients (5 males aged between 3 and 22; and 17 females aged between 1.5 and 19) and total of 24 recordings with 256 Hz sampling frequency. In our study, 10-channel (FP1-F7, F7-T7, T7-P7, FP1-F3, T7-FT9, FP2-F8, F8-T8, T8-P8, FP2-F4, FT10-T8) EEG signals of the first 23 subjects are utilized to satisfy the similar channel positions with the IKCU dataset. On this dataset, the proposed approach is applied for (i) segment-based seizure detection: using 10-min inter-seizure segments prior to minimum of 1 h from onset of all selected seizures were extracted considering the data description of the CHB-MIT dataset and seizure segments, (ii) segment-based seizure prediction: using 5-min pre-seizure EEG segments just before seizure onset and inter-seizure EEG segments. 10-min inter-seizure segments prior to the minimum of 1 h to the onset of the selected all seizure segments of a patient were obtained considering data description of the CHB-MIT dataset.

3. Results

In this study, the proposed method based on the TF analysis, described in Sec. 2, is tested on IKCU and CHB-MIT datasets. All training and testing procedures were performed on Nvidia GeForce RTX 2080 TI graphics processing unit and 64 GB random-access memory using Tensor Flow 2.1 and Cuda 10.2.

Two different cross-validation methods are adopted to evaluate the robustness of the proposed models. The k-fold cross-validation (CV)[24] method is utilized based on a patient-correlated detection (PCD) approach where the trained model contains a mixed amount of data from all patients. Hence, the trained model preserves some of the characteristics of all patients used for training. Although this method does not evaluate well the accuracy in testing new patients, it is more successful in testing the data unused for training the model. The other method used for testing our results is the leave-one-out cross-validation (LOOCV).[38] In the LOOCV method, the subjects-folds used for testing are not considered for training. Therefore, it can be used for patient-independent detection (PID) validation. In the PID method, since the patient data used as test data are not included in any training phase, the trained model does not have any characteristic features of these patient data. Therefore, compared to the PCD method, this method better evaluates accuracy in testing new patients. Moreover, the LOOCV method is advantageous in choosing a more robust model compared to other CV methods.[39] In this study, since patients are selected and excluded, the LOOCV method works as a leave-subject-out CV approach.

Additionally, recall (REC), precision (PRE), accuracy (ACC), F1-Score (F1-S), area under the receiver operating characteristic curve (ROC-AUC),[4,8,19] and mean squared error (MSE)[38] values are calculated during the segment-based evaluation phase[40] of the proposed models, in order to

investigate the imbalanced dataset effects further. Moreover, imbalance ratio (IR)[41] values are calculated for all training sets.

3.1. *Experimental results and discussion using IKCU dataset*

Using the IKCU dataset which includes 10-channel EEG signals of 16 epilepsy patients, *segment-based seizure detection* approach is performed. All pre-seizure and seizure EEG segments are divided into nonoverlapping 1 s- and 5 s-long segments. The high-resolution TFR of each EEG segment is calculated using the SST method. The magnitude square of SST matrices are considered as images and used as input for the proposed CNN architecture. PID and PCD validation methods are used to test the performance of the proposed image-based CNN approach. Two different segment durations are tested in order to reveal the contribution of the duration–TF resolution to the performance of the proposed approach. The number of segments used to training and test phases of the proposed approach for all cases are given in Table 1.

In the PCD approach, IR is 1 for both 1 s and 5 s long segments. In the PID approach, IR is calculated as 1.004 and 1.011 for 1 s and 5 s segments, respectively. Hence, no imbalanced dataset effect was encountered during the training of the IKCU dataset. Initially, we compared the proposed ResNet-50-based architecture to the traditional ResNet-50 architecture. The comparison was conducted with the 5 s long SST segments obtained from the IKCU dataset. Traditional ResNet-50 is tested using PCD and PID validation methods, and average ACC of 94.78% and 90.97% were obtained respectively. The proposed architecture achieved an average ACC of 99.06% and 97.22% and revealed an average of 5.27% significant difference compared to the traditional ResNet-50 architecture. Further, it takes 519 s to train the traditional ResNet-50 architecture and 428 s to train the proposed modified ResNet-50 architecture using 5 s long SST images. Thus, the proposed architecture provides a 17.53% gain of training time. Therefore, all subsequent experiments are conducted on the proposed architecture. Additionally, we compare our results by the SST approach with spectrograms obtained by the traditional TFR approach, STFT. In our experiments, a Hamming window and overlap of 50% are utilized for STFT calculation for both 1 s and 5 s long EEG segments. The seizure detection performance for all cases in terms of REC, PRE, ACC, cross-entropy (CE)-Loss, and F1-S for the segment-based analysis are presented in Tables 2–5.

Table 1. The number of segments reserved for training and testing for IKCU dataset.

	1 s Duration		5 s Duration	
CV Type	5-fold-CV	LOOCV	5-fold-CV	LOOCV
Tr-S	7856	9206	1520	1781
Te-S	1964	614	380	119
Tot-S	**9820**	**9820**	**1900**	**1900**

Notes: Tr-S = Training, Te-S = Testing, and Tot-S = Total Size.

Segment-based seizure detection performance for the 1 s segment duration using PID method is presented in Table 2. For 1 s segment duration, while by using the SST-based CNN approach, the highest patient-independent seizure detection performance (97.72% ACC, 100% PRE, 95.40% REC, and 97.62% F1-S) is achieved for "Patient-15", the maximum seizure detection performance (89.57% ACC, 91.03% PRE, 87.41% REC, and 89.18% F1-S) is obtained utilizing the STFT-based CNN approach for "Patient-16". For the same validation model of 5 s segment duration, similarly, maximum seizure detection performance (100% ACC, PRE, REC, and F1-S) is achieved using the SST-based CNN approach for "Patient-14". However, lower seizure detection performances have been achieved using STFT-based CNN approach (shown in Table 3). In addition to the performance evaluation values given in Tables 2 and 3, and the average ROC-AUC values are calculated. In the PID approach, the ROC-AUC values obtained using SST 1 s, STFT 1 s, SST 5 s, and STFT 5 s are 0.963, 0.912, 0.987, and 0.944, respectively.

We also tested the performance of SST- and STFT-based CNN approaches for segment-based epileptic seizures detection using PCD validation method (Table 4). For 1 s segment duration, the best seizure detection accuracy calculated using the "Fold-2" PCD model of SST-based CNN approach is 96.54%, which is significantly higher than the best accuracy calculated using the "Fold-1" PCD model of STFT-based CNN approach (88.34%). Similarly,

Table 2. The segment-based seizure detection performance obtained using the PID validation model for 1 s long EEG segments of the IKCU dataset.

	SST 1 s (%)				STFT 1 s (%)			
Patient	ACC	PRE	REC	F1-S	ACC	PRE	REC	F1-S
Patient-1	96.58	98.72	94.77	96.70	86.97	89.74	85.37	87.50
Patient-2	96.09	98.40	94.17	96.24	86.16	88.78	84.71	86.70
Patient-3	93.65	97.12	90.99	93.96	82.74	85.26	81.60	83.39
Patient-4	91.37	96.47	87.76	91.91	82.74	84.94	81.79	83.34
Patient-5	91.21	96.79	87.28	91.79	83.39	85.90	82.21	84.01
Patient-6	95.77	98.08	93.87	95.93	85.69	88.18	84.40	86.25
Patient-7	95.77	98.40	93.60	95.94	85.18	87.82	83.79	85.76
Patient-8	96.74	99.04	94.79	96.87	87.79	90.38	86.24	88.26
Patient-9	95.90	98.40	93.88	96.09	87.79	90.71	86.02	88.30
Patient-10	95.28	97.76	93.27	95.46	85.18	87.50	84.00	85,71
Patient-11	94.95	97.44	92.97	95.15	85.50	88.14	84.10	86.07
Patient-12	91.69	96.79	88.05	92.21	83.50	85.90	82.46	84.14
Patient-13	92.02	97.12	88.34	92.52	82.90	85.26	81.85	83.52
Patient-14	97.39	99.68	95.40	97.49	88.27	91.03	86.59	88.75
Patient-15	**97.72**	100.0	95.36	97.62	88.76	91.67	86.93	89.24
Patient-16	97.07	99.36	95.09	97.18	**89.57**	91.03	87.41	89.18
Average	**94.95**	**98.10**	**92.47**	**95.19**	**85.76**	**88.27**	**84.34**	**86.26**

Table 3. The segment-based seizure detection performance obtained using the PID validation model for 5 s long EEG segments of the IKCU dataset.

	SST 5 s (%)				STFT 5 s (%)			
Patient	ACC	PRE	REC	F1-S	ACC	PRE	REC	F1-S
Patient-1	99.16	100.0	98.39	99.19	96.64	96.72	96.72	96.72
Patient-2	98.32	100.0	96.83	98.39	94.96	95.08	95.08	95.08
Patient-3	96.64	96.72	96.72	96.72	90.76	93.44	89.06	91.20
Patient-4	94.96	95.08	95.08	95.08	89.92	91.80	88.89	90.32
Patient-5	94.12	95.08	93.55	94.31	89.92	90.16	90.16	90.16
Patient-6	97.48	96.72	98.33	97.52	93.28	93.44	93.44	93.44
Patient-7	97.48	98.36	96.77	97.56	95.80	96.72	95.16	95.93
Patient-8	99.16	100.0	98.39	99.19	97.48	98.36	96.77	97.56
Patient-9	99.16	98.36	100.0	99.17	**99.16**	100.0	98.27	99.13
Patient-10	98.32	98.36	98.36	98.36	95.80	95.80	96.67	96.23
Patient-11	98.32	98.36	98.36	98.36	94.12	95.08	93.55	94.31
Patient-12	92.44	93.44	91.94	92.68	90.76	95.08	87.88	91.34
Patient-13	91.60	93.44	90.48	91.94	89.92	95.08	86.57	90.63
Patient-14	**100.0**	100.0	100.0	100.0	98.32	98.36	98.36	98.36
Patient-15	99.16	100.0	98.39	99.19	97.48	98.36	96.77	97.56
Patient-16	99.16	98.36	100.0	99.17	96.64	96.72	96.72	96.72
Average	**97.22**	**97.64**	**96.97**	**97.30**	**94.44**	**95.64**	**93.75**	**94.67**

for the 5 s segment duration, the highest seizure detection accuracy (99.47%) for the SST-based CNN approach is achieved using "Fold-4" PCD model, but the maximum seizure detection accuracy (98.68%) is obtained using the "Fold-2" PCD model for the STFT-based CNN approach. Additionally, for both 1 s and 5 s long segment durations, all average recall, precision, and F1-score values calculated using the SST-based CNN approach (for 1 s: 93.56% REC, 97.01% PRE, and 95.25% F1-S; and for 5 s: 99.18% REC, 98.99% PRE, and 99.08% F1-S) are significantly higher than those calculated for the STFT-based CNN approach (for 1 s: 86.88% REC, 88.56% PRE, and 87.71% F1-S; for 5 s: 95.37% REC, 98.01% PRE, and 96.67% F1-S). It is also observed that the average CE-Loss calculated using the SST-based CNN approach is lower than that of the STFT-based CNN approach for all cases. In the PCD approach,

Table 4. The segment-based seizure detection performance obtained using the PCD validation model for both 1 s and 5 s long EEG segments of the IKCU dataset.

		SST					STFT				
	Folds	ACC	Loss	PRE	REC	F1-S	ACC	Loss	PRE	REC	F1-S
1 s	Fold-1	95.42	0.256	97.03	94.23	95.61	**88.34**	0.387	88.77	87.00	87.88
	Fold-2	**96.54**	0.242	98.68	94.13	96.35	85.92	0.450	87.75	85.30	86.51
	Fold-3	93.78	0.289	96.26	91.96	94.06	86.83	0.423	87.75	86.80	87.27
	Fold-4	96.08	0.243	96.94	95.53	96.23	88.03	0.396	89.31	87.65	88.47
	Fold-5	93.71	0.289	96.14	91.95	94.00	87.79	0.398	89.22	87.66	88.43
	Average	**95.11**	**0.264**	**97.01**	**93.56**	**95.25**	**87.38**	**0.411**	**88.56**	**86.88**	**87.71**
5 s	Fold-1	99.25	0.017	98.51	100.0	99.25	94.21	0.121	96.52	92.82	94.63
	Fold-2	98.95	0.018	99.00	99.00	99.00	**98.68**	0.075	100.0	97.57	98.77
	Fold-3	99.21	0.018	99.44	99.01	99.22	96.32	0.086	98.01	95.17	96.57
	Fold-4	**99.47**	0.017	100.0	98.88	99.44	95.53	0.098	97.51	94.23	95.84
	Fold-5	98.42	0.019	98.01	98.99	98.50	97.37	0.079	98.01	97.04	97.52
	Average	**99.06**	**0.018**	**98.99**	**99.18**	**99.08**	**96.42**	**0.092**	**98.01**	**95.37**	**96.67**

Notes: Loss: Cross-Entropy Loss; all ACC, REC, PRE, and F1-S values are given as %.

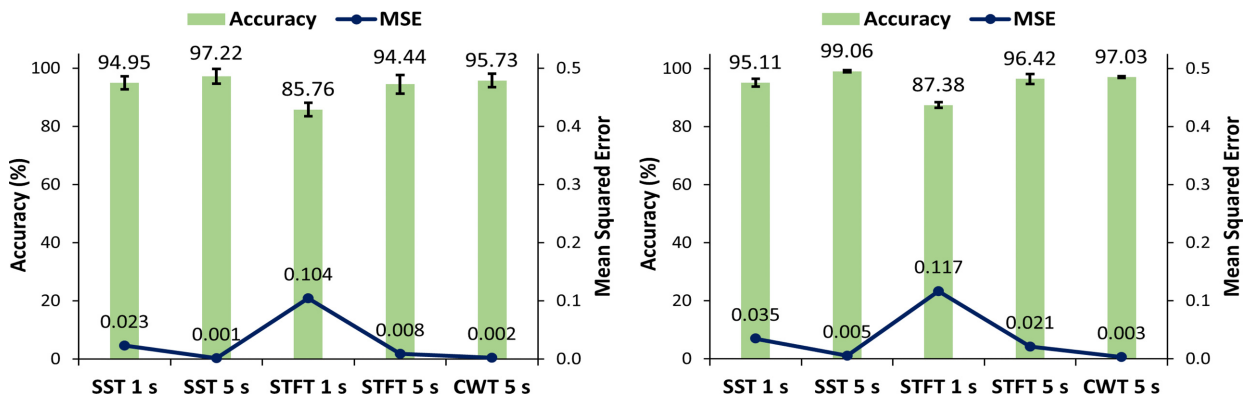


Fig. 4. Comparison of average accuracy and MSE obtained using (a) PID (standard deviations for accuracy values: ∓2.24% (SST 1 s), ∓2.58% (SST 5 s), ∓2.27% (STFT 1 s), ∓3.26% (STFT 5 s), and ∓2.31% (CWT 5 s)) and (b) PCD (standard deviations for accuracy values: ∓1.30% (SST 1 s), ∓0.40% (SST 5 s), ∓0.99% (STFT 1 s), ∓1.71% (STFT 5 s), and ∓0.35% (CWT 5 s))-based validation models for IKCU dataset.

the average ROC-AUC values obtained using SST 1 s, STFT 1 s, SST 5 s, and STFT 5 s are 0.970, 0.942, 0.996, and 0.973, respectively.

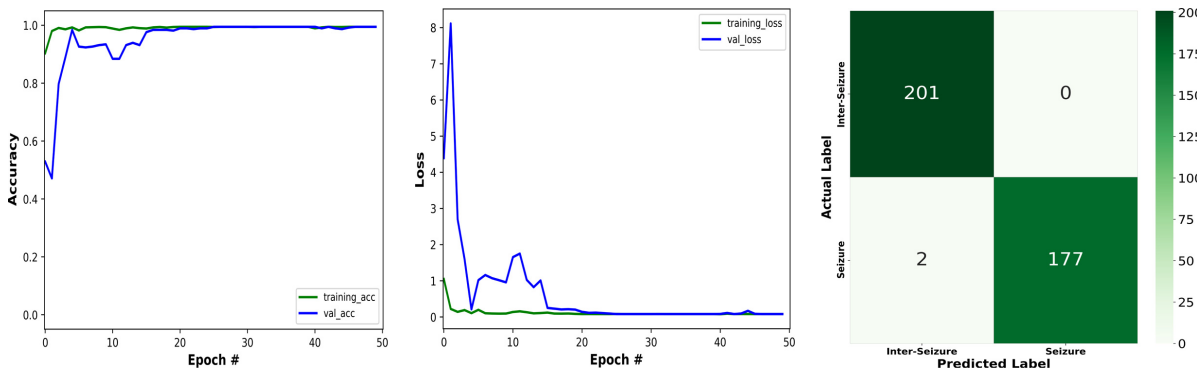
In addition, to further explore the advantages of the proposed SST-based CNN approach, all steps of the classification algorithm are repeated utilizing a CWT on 5 s long segments. In our experiments, Morlet wavelet is employed, a frequently utilized mother wavelet, to calculate the scalogram of EEG segments,[42] which are used to train the network. The average accuracy and MSE values are demonstrated for tested TFR methods, and validation models in Figs. 4(a) and 4(b). It is noteworthy that, while the CWT-based approach yields higher average ACC and lower average MSE values compared to the STFT method, the proposed SST approach provided the highest average ACC and lowest average MSE values for both validation models. Additionally, each approach provides higher average performance metrics using PCD-based validation model.

As shown in Table 5, resulting ACC and calculated training durations are given for various learning rate hyperparameters in the PCD method. The best scenario in all segments occurred at a 0.001 learning rate. Although the models provide lower training time at higher learning rates, they reached local minima points. Using a lower learning rate of 0.001, we reach the global minimum. The best training and validation ACC, CE-Loss values, and the best confusion matrix in the test phase of the PCD method

Table 5. Average accuracy and training duration comparison versus learning rate hyperparameters in PCD-based approach for each fold in the IKCU dataset.

	SST 1 s		SST 5 s		STFT 1 s		STFT 5 s	
Learning rate	ACC (%)	TT (s)	ACC (%)	TT (s)	ACC (%)	TT (s)	ACC (%)	TT (s)
0.1	92.45	578	96.95	111	83.69	613	94.68	98
0.01	93.55	1093	98.41	234	84.96	1058	95.89	189
0.001	**95.11**	**2168**	**99.06**	**428**	**87.38**	**2000**	**96.42**	**346**
0.0001	94.67	4076	97.43	973	86.98	4004	95.76	798

Notes: TT = Training Time.



(a) (b) (c)

Fig. 5. Example performance graphs obtained in the 4th-fold of PCD method with 5 s long SST segments; the best (a) training and validation accuracies, (b) CE-Losses, and (c) the confusion matrix in the IKCU dataset.

with 5 s SST segments are given in Fig. 5. Note that, ACC values converge to the upper and CE-Loss values converge to the lower limits. Therefore, overfitting or underfitting are not observed in any of the trained models. Notice also that SST 5 s models are converged before 50 epochs, hence it is fixed to 50 to compare all models on equal conditions. Moreover, SST 5s model is almost fully capable of distinguishing seizure from pre-seizure.

Based on the conducted experiments, the highest performing 5 s long SST images are selected and used to train several well-known architectures. Testing phase performance comparisons with that of the proposed network are shown in Fig. 6. Results emphasize the superiority of the proposed architecture in classifying SST images. Notice that, AlexNet yields the lowest classification performance. The proposed ResNet-based architecture provides a significant improvement of 5.27% compared to traditional ResNet-50,[32,35] 24.33% compared to AlexNet,[34,43] 12.86% compared to VGG16,[35] and 15.59% compared to SqueezeNet.[34]

3.2. *Experimental results and discussion using CHB-MIT scalp EEG database*

We applied the proposed SST-based CNN approach to CHB-MIT Epileptic EEG dataset which was previously employed in other DL studies, and compared our performance with existing literature. Two different classification problems are addressed using this dataset. *Segment-based seizure detection*: the SST-based CNN model is trained to differentiate between inter-seizure and seizure EEG segments, *Segment-based seizure prediction*: the model is trained to distinguish inter-seizure and pre-seizure EEG segments. Signals in this dataset are sampled with 256 Hz providing a sufficient TF resolution. Hence, the proposed approach is only conducted using 1 s long segment duration. Using LOOCV- and CV-based validation models for segment-based seizure detection and prediction, we evaluate the performance of our proposed approach.

In addition, the patient-independent prediction (PIP) approach similar to the PID method, and the

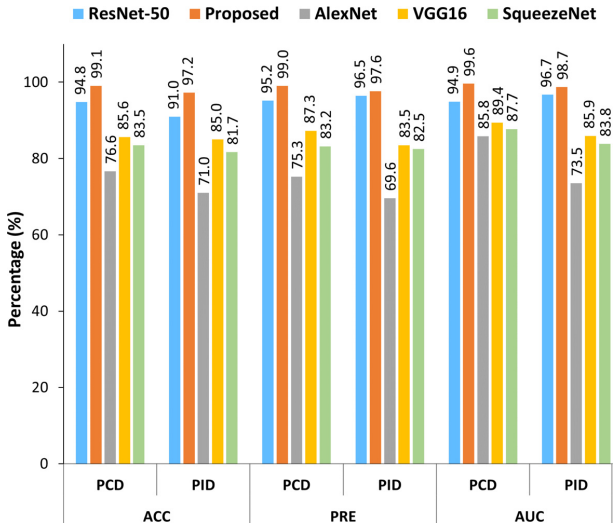


Fig. 6. Comparison of proposed architecture with well-known CNN architectures using 5 s long SST segment in IKCU dataset.

patient-correlated prediction (PCP) approach similar to the PCD method utilized above were used to validate the success of segment-based seizure prediction models on the CHB-MIT dataset. For this experiment, variable number of training and testing sets are formed because the number of seizure segments of each patient is different. Using the CHB-MIT dataset, in the PCD-based method, a total of 23,080 SST segments are used, with 18,464 for training and the rest of 4616 for testing in each fold. This data includes a total of 9280 seizure segments and a total of 13,800 inter-seizure segments (600 segments for each subject). Seizure segments are variable in each patient and contain an average of 403 (min 66 and max 1396) SST 1 s segments. In this method, the IR was calculated as 1.49. In the PCP-based method, a total of 20,700 SST segments are used, where 16,560 segments used for training and 4140 for testing at each fold. These segments include a total of 6900 pre-seizure segments (300 segments for each subject) and a total of 13,800 inter-seizure segments (600 segments for each subject). Furthermore, in the PIP-based method, 19,800 segments are used for training and 900 segments for testing at each fold. In both PCP and PIP methods, the IR value was calculated as 2. As the PID method contains a variable number of seizure segments like the PCD method, an average of 22,077 segments for training and 1003 segments for testing are used in each fold. In this method, the IR value was calculated as 2.25. The reason for the data imbalance is the variation in the number of seizure segments in CHB-MIT database. Although IR affects our models, its level is negligible. Performance evaluation results of the experiments carried out with the proposed approach are demonstrated in Tables 6 and 7 for the segment-based seizure detection and prediction tasks, respectively.

It can be observed from Table 6 that generally high segment-based seizure detection performances in terms of accuracy ($\geq$ 99.35%), precision ($\geq$ 99.64%), recall ($\geq$ 99.10%), and F1-score ($\geq$ 99.46%) are achieved using the PCD validation model. The proposed SST-based CNN approach provides equally high performance for the more challenging segment-based seizure prediction problem with accuracy ($\geq$ 96.33%), precision ($\geq$ 98.88%), recall ($\geq$ 95.75%), and F1-score ($\geq$ 97.29%) for the PCP validation model. On the other hand, by using the PID validation model (given in Table 7), while maximum segment-based seizure detection performances (100% ACC, PRE, REC, and F1-S) are

Table 6. The segment-based seizure detection (inter-seizure versus seizure) and segment-based seizure prediction (inter-seizure versus pre-seizure) performances of the proposed SST-based CNN approach obtained using the CV-based validation model for 1 s EEG segments of the CHB-MIT dataset.

	Seizure detection					Seizure prediction				
Folds	ACC	Loss	PRE	REC	F1-S	ACC	Loss	PRE	REC	F1-S
Fold-1	99.76	0.011	99.86	99.75	99.80	96.81	0.143	99.13	96.20	97.64
Fold-2	99.46	0.015	99.64	99.46	99.55	97.10	0.115	99.28	96.48	97.86
Fold-3	99.61	0.013	99.71	99.64	99.67	**97.71**	0.087	99.46	97.17	98.30
Fold-4	99.35	0.017	99.82	99.10	99.46	96.33	0.194	98.88	95.75	97.29
Fold-5	**99.98**	0.009	100.0	99.66	99.83	96.62	0.169	98.99	96.06	97.50
Average	**99.63**	**0.013**	**99.81**	**99.52**	**99.66**	**96.91**	**0.142**	**99.15**	**96.33**	**97.72**

Notes: Loss: Cross-Entropy Loss; all ACC, REC, PRE, and F1-S values are given as %.

Table 7. The segment-based seizure detection (inter-seizure versus seizure) and segment-based seizure prediction (inter-seizure versus pre-seizure) performances of the proposed SST-based CNN approach obtained using the LOOCV-based validation model for 1 s EEG segments of the CHB-MIT dataset.

	Seizure detection (%)				Seizure prediction (%)			
Patient	ACC	PRE	REC	F1-S	ACC	PRE	REC	F1-S
chb01	99.17	99.33	99.17	99.25	99.44	99.67	99.50	99.58
chb02	99.09	99.16	99.66	99.41	98.88	99.17	99.17	99.17
chb03	98.60	99.67	99.00	99.33	98.65	99.00	99.00	99.00
chb04	98.96	98.83	99.66	99.24	99.22	99.50	99.33	99.41
chb05	99.22	99.67	98.84	99.25	98.31	98.67	98.83	98.75
chb06	99.59	100.0	99.50	99.75	97.75	98.33	98.33	98.33
chb07	**100.0**	100.0	100.0	100.0	99.78	99.83	99.83	99.83
chb08	99.34	99.83	98.52	99.17	96.63	97.33	97.66	97.49
chb09	99.89	99.83	100.0	99.91	**100.0**	100.0	100.0	100.0
chb10	99.14	99.67	98.84	99.25	96.18	97.00	97.32	97.16
chb11	99.50	99.17	99.66	99.41	98.78	99.17	99.00	99.08
chb12	99.68	99.33	100.0	99.66	95.96	96.83	97.16	96.99
chb13	99.30	99.50	99.50	99.50	94.44	96.67	95.08	95.87
chb14	99.48	99.50	99.83	99.66	99.11	99.50	99.17	99.33
chb15	99.60	99.67	99.01	99.34	96.18	97.00	97.32	97.16
chb16	98.06	99.17	98.67	98.92	95.73	96.67	96.99	96.83
chb17	98.66	99.33	98.68	99.00	95.56	97.50	95.90	96.69
chb18	98.26	98.50	98.83	98.66	99.00	99.33	99.17	99.25
chb19	99.16	99.50	99.33	99.41	99.33	99.67	99.34	99.50
chb20	99.66	99.83	99.67	99.75	**100.0**	100.0	100.0	100.0
chb21	97.00	98.67	97.37	98.02	93.89	96.33	94.60	95.46
chb22	99.50	100.0	99.34	99.67	99.67	99.83	99.67	99.75
chb23	**100.0**	100.0	100.0	100.0	99.56	99.50	99.83	99.66
Average	**99.17**	**99.49**	**99.26**	**99.37**	**97.92**	**98.54**	**98.36**	**98.45**

obtained for patients chb07 and chb23, the highest segment-based seizure prediction performances are achieved (100% ACC, PRE, REC, and F1-S) for patients chb09 and chb20 using PIP approach. Additionally, the average accuracy and MSE values are given in terms of conducted task and validation models, in Fig. 7. Note that the PCD validation model shows higher average segment-based seizure detection ACC (99.63%) and lower average MSE values (0.003), PIP validation model provides better performance (97.92% ACC, and 0.02 MSE) for segment-based seizure prediction task.

Performance comparison of recent segment-based seizure detection and prediction studies conducted on the CHB-MIT dataset with the proposed approach is demonstrated in Table 8. The best performance evaluation results are highlighted in boldface numbers. Hossain *et al.*[44] introduced a segment-based seizure detection approach in which the multi-channel raw EEG data were used as input to the CNN model. Using the cross patient validation

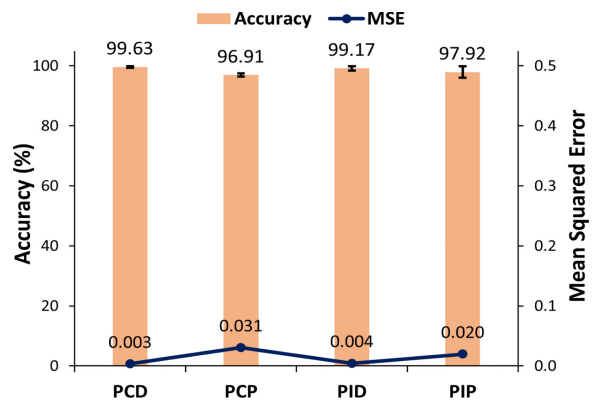


Fig. 7. Comparison of average accuracy and MSE values obtained using segment-based PCD, PCP, PID, and PIP models (standard deviations for accuracy values: ∓0.25%, ∓0.53%, ∓0.69%, and ∓1.90%, respectively) in the CHB-MIT dataset.

model, the proposed approach provides the average 90.00% SEN, 91.65% SPE, 98.05% ACC for all patients. In another study,[43] four-class segment-based seizure detection approach was proposed.

Table 8. Comparison of recent segment-based seizure detection and prediction studies conducted using CHB-MIT dataset with proposed work.

Approach	Study	No. of patients	Feature	Classifier	No. of seizures	ACC (%)	REC (%)
Seizure detection	Ayodele *et al.*[28]	23	Features maps	RCNN	198	—	71.45
	Li *et al.*[24]	24	EEG time-series	NLSTM	127	95.29	95.42
	Liang *et al.*[43]	23	EEG waveforms	LRCN	198	99	84
	Hossain*et al.*[44]	23	EEG time-series	CNN	198	98.05	90.00
	Thodorof *et al.*[27]	23	Image based	LSTM	198	—	85.00
	Wei *et al.*[45]	24	EEG time-series	CNN	151	81.49	70.68
	Zhang *et al.*[35]	9	STFT	CNN	—	98.26	98.01
	Zhou *et al.*[46]	24	EEG time series	CNN	—	62.3	61.2
	Zhou *et al.*[46]	24	EEG frequency-series	CNN	—	97.5	96.9
	This work[a]	**23**	SST	CNN	**144**	**99.17**	**99.26**
	This work[b]	**23**	SST	CNN	**144**	**99.63**	**99.52**
Seizure prediction	Liu *et al.*[47]	2	EEG features	CNN	36	—	91.5
	Shahbazi *et al.*[48]	14	STFT	LSTM	49	—	98.21
	Tsiouris *et al.*[40]	24	EEG features	LSTM	185	—	99.84
	Zhou *et al.*[46]	24	EEG time series	CNN	—	59.5	61.8
	Zhou *et al.*[46]	24	EEG frequency-series	CNN	—	95.6	94.2
	This work[a]	**23**	SST	CNN	**144**	**97.92**	**98.36**
	This work[b]	**23**	SST	CNN	**144**	**99.91**	**96.33**

Notes: [a]Performance evaluation results obtained using PID-based validation model.
[b]Performance evaluation results obtained using PCD-based validation model.
REC: Sensitivity (SEN).

EEG waveform images of normal, pre-seizure, seizure and post-seizure EEG data were used as input to an 18-layer Long-Term recurrent convolutional network (LRCN). By using cross patient seizure detection model, 84% sensitivity, 99% specificity, and 99% accuracy were achieved. The segment-based seizure detection approach was proposed by using the multi-channel time-series EEG signals-based 12-layer CNN model in another study. The merger of the increasing and decreasing sequences (MIDS) and data augmentation method was utilized, separately, and achieved cross patient performance evaluation results. Using multi-channel EEG, 70.68% sensitivity, 81.49% ACC, and 92.30% specificity values were achieved.[45] In a study conducted by Zhang *et al.*,[35] three different CNN models (VGG16, VGG19, and ResNet-50) based segment-based seizure detection approaches were performed using STFT-based preprocessing step. Maximum 98.26% seizure detection accuracy was achieved using nine patients' EEG recordings. It is noteworthy that among all seizure detection methods compared, both PID and PCD validation results of our SST-based CNN methods are significantly higher.

Further, the segment-based seizure prediction performance of the proposed SST-based CNN approach is tested and compared with similar studies in the literature (given in Table 8). The 3-layer CNN-based seizure detection and prediction method using the CHB-MIT dataset has been presented using both time domain and frequency domain signals as inputs for classification.[46] In the segment-based seizure detection, while this approach presented an average of 62.3% ACC and 61.2% REC values for the time domain inputs, for the frequency domain inputs average of 97.5% ACC and 96.9% REC values were presented. Additionally, in the segment-based seizure prediction, average of 95.6% ACC and 94.2% REC values were achieved for the frequency domain inputs. LSTM-based CNN method is utilized in another segment-based seizure prediction study. Spectrograms of corresponding signals were

used as input. For 14 patients, an average of 98.21 SEN was obtained.[48] The performance of segment-based seizure prediction obtained in that study is higher than that of our study. While EEG data of 14 patients and 49 seizure segments are examined in that study, EEG data of 23 patients and 144 seizure segments are evaluated in our study. In another segment-based seizure prediction study,[40] statistical moments, zero crossings, WT coefficients, PSDs of EEG subbands, and cross-correlation were calculated as features and utilized as input for LSTM. For the pre-seizure window of 120 min, an average of 99.84% SEN and 99.86% SPE were achieved for 24 patients. In that study, higher segment-based seizure prediction sensitivity was obtained compared to the proposed approach at the expense of comprehensive and computationally expensive feature extraction step.

Seizure events cause changes in time and frequency characteristics of EEG signals. Thus, the correct distribution of epileptic EEG signals energy into the joint time-frequency plane is very effective in the classification process. SST is a recently developed TF reassignment method that provides a close to an ideal representation of nonstationary signals in the TF plane. In our recent work,[49] TF representations obtained by SST are used as input for both machine and DL methods using only IKCU Epilepsy dataset. Higher order TF moments are manually extracted and classified using three classifiers. Higher classification performance is achieved by using a standard CNN architecture trained by SST images. The present study presents utilization of the high-resolution TF representations obtained by SST in the segment-based classification of epileptic EEG signals. We propose treating those SST energy distributions as images and utilize them in the training of a deep network. Among other network architectures, the proposed modified ResNet-50 architecture yielded better classification performance. The layer-filters of the proposed architecture are optimized according to our classification problem. Further, the hyperparameters used in the training of the network are fine-tuned by making performance comparisons. While constructing the deep network, it was aimed to extract all valuable information of the energy distribution in TFR images with residual features. Experimental results have emphasized the advantages of the proposed architecture in the segment-based classification problem.

4. Conclusion

In this study, a novel and effective approach is proposed to detect and predict epileptic seizures using EEG signals. TF spectra of EEG segments are obtained by SST transform, which is a recent and popular high-resolution TF analysis approach. These TF images are used to train a ResNet-based CNN architecture. The proposed method was tested on two different datasets (IKCU and CHB-MIT) in order to validate its robustness, and compare its success with existing studies. Moreover, two different CV methods have been used to prove the transparency of the obtained results. Especially, SST was used the first time to analyze epileptic EEG signals in conjunction with DL, as well as TF-images obtained from SST were used to train a CNN architecture the first time, to the best of our knowledge. Results demonstrated that outstanding performance is achieved in predicting and detecting seizures from EEG signals. Particularly, the validation of the proposed models with the PIP-based method indicates that it works well in testing new patients. The proposed SST-based method having lower computational complexity $(\mathcal{O}(N_v N \log^2 N))$[11] than most of the high-resolution TF approaches,[8,11,12] and encouraging performance evaluation results suggest that the proposed DL approach may be used especially in pre-seizure alert applications.

References

1. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy, *IEEE Trans. Biomed. Eng.* **54**(2) (2007) 205–211.
2. U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan and H. Adeli, Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals, *Comput. Biol. Med.* **100** (2018) 270–278.
3. H. Adeli, Z. Zhou and N. Dadmehr, Analysis of EEG records in an epileptic patient using wavelet transform, *J. Neurosci. Methods* **123**(1) (2003) 69–87.
4. J. Xiang, C. Li, H. Li, R. Cao, B. Wang, X. Han and J. Chen, The detection of epileptic seizure signals based on fuzzy entropy, *J. Neurosci. Methods* **243** (2015) 18–25.
5. U. R. Acharya, F. Molinari, S. V. Sree, S. Chattopadhyay, K.-H. Ng and J. S. Suri, Automated diagnosis of epileptic EEG using entropies, *Biomed. Signal Process. Control* **7**(4) (2012) 401–408.

6. P. P. M. Shanir, K. A. Khan, Y. U. Khan, O. Farooq and H. Adeli, 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.
7. E. Alickovic, J. Kevric and A. Subasi, Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction, *Biomed. Signal Process. Control* **39** (2018) 94–102.
8. O. K. Cura, S. K. Atli, H. S. Türe and A. Akan, Epileptic seizure classifications using empirical mode decomposition and its derivative, *BioMed. Eng. OnLine* **19**(1) (2020) 1–22.
9. A. Zahra, N. Kanwal, N. ur Rehman, S. Ehsan and K. D. McDonald-Maier, Seizure detection from EEG signals using multivariate empirical mode decomposition, *Comput. Biol. Med.* **88** (2017) 132–141.
10. A. Ahrabian, D. Looney, L. Stanković and D. P. Mandic, Synchrosqueezing-based time-frequency analysis of multivariate data, *Signal Process.* **106** (2015) 331–341.
11. S. Mamli and H. Kalbkhani, Gray-level co-occurrence matrix of fourier synchro-squeezed transform for epileptic seizure detection, *Biocybern. Biomed. Eng.* **39**(1) (2019) 87–99.
12. N. A. Khan and S. Ali, A new feature for the classification of non-stationary signals based on the direction of signal energy in the time–frequency domain, *Comput. Biol. Med.* **100** (2018) 10–16.
13. L. Boubchir, S. Al-Maadeed and A. Bouridane, On the use of time-frequency features for detecting and classifying epileptic seizure activities in non-stationary EEG signals, 2014 *IEEE Int. Conf. Acoustics, Speech and Signal Processing* (*ICASSP*) (2014), pp. 5889–5893.
14. S. Ghosh-Dastidar, H. Adeli and N. Dadmehr, Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection, *IEEE Trans. Biomed. Eng.* **54**(9) (2007) 1545–1551.
15. A. T. Tzallas, M. G. Tsipouras and D. I. Fotiadis, Automatic seizure detection based on time-frequency analysis and artificial neural networks, *Comput. Intell. Neurosci.* **2007** (2007).
16. M. Sharma, R. B. Pachori and U. R. Acharya, A new approach to characterize epileptic seizures using analytic time-frequency flexible wavelet transform and fractal dimension, *Pattern Recogn. Lett.* **94** (2017) 172–179.
17. Y. Li, W. Cui, M. Luo, K. Li and L. Wang, Epileptic seizure detection based on time-frequency images of EEG signals using gaussian mixture model and gray level co-occurrence matrix features, *Int. J. Neural Syst.* **28**(07) (2018) p. 1850003.
18. D. Chen, S. Wan, J. Xiang and F. S. Bao, A high-performance seizure detection algorithm based on discrete wavelet transform (DWT) and EEG, *PloS ONE* **12**(3) (2017) p. e0173138.
19. O. K. Cura and A. Akan, Classification of epileptic EEG signals using synchrosqueezing transform and machine learning, *Int. J. Neural Syst.* (2020).
20. Q. Yuan, W. Zhou, L. Zhang, F. Zhang, F. Xu, Y. Leng, D. Wei and M. Chen, Epileptic seizure detection based on imbalanced classification and wavelet packet transform, *Seizure* **50** (2017) 99–108.
21. A. H. Ansari, P. J. Cherian, A. Caicedo, G. Naulaers, M. De Vos and S. Van Huffel, Neonatal seizure detection using deep convolutional neural networks, *Int. J. Neural Syst.* **29**(4) (2019) 1850011.
22. L.-C. Lin, C.-S. Ouyang, R.-C. Wu, R.-C. Yang and C.-T. Chiang, Alternative diagnosis of epilepsy in children without epileptiform discharges using deep convolutional neural networks, *Int. J. Neural Syst.* **30**(5) (2020) 1850060.
23. Z. Ma, Reachability analysis of neural masses and seizure control based on combination convolutional neural network, *Int. J. Neural Syst.* **30**(1) (2020) 1950023.
24. Y. Li, Z. Yu, Y. Chen, C. Yang, B. Li *et al.*, Automatic seizure detection using fully convolutional nested lstm., *Int. J. Neural Syst.* **30**(4) (2020) 2050019–2050019.
25. J. Thomas, J. Jin, P. Thangavel, E. Bagheri, R. Yuvaraj, J. Dauwels, R. Rathakrishnan, J. J. Halford, S. S. Cash and B. Westover, Automated detection of interictal epileptiform discharges from scalp electroencephalograms by convolutional neural networks, *Int. J. Neural Syst.* **30**(11) (2020) 2050030.
26. G. Liu, W. Zhou and M. Geng, Automatic seizure detection based on s-transform and deep convolutional neural network, *Int. J. Neural Syst.* **30**(4) (2020) 1950024.
27. P. Thodoroff, J. Pineau and A. Lim, Learning robust features using deep learning for automatic seizure detection, in *Proc. 1st Machine Learning for Healthcare Conf.*, Vol. 56 (PMLR, Northeastern University, Boston, MA, USA), 18–19 August 2016, pp. 178–190.
28. K. Ayodele, W. Ikezogwo, M. Komolafe and P. Ogunbona, Supervised domain generalization for integration of disparate scalp EEG datasets for automatic epileptic seizure detection, *Comput. Biol. Med.* **120** (2020) 103757.
29. H. S. Nogay and H. Adeli, Detection of epileptic seizure using pretrained deep convolutional neural network and transfer learning, *Eur. Neurol.* **83**(6) (2020) 602–614.
30. G. Yu, M. Yu and C. Xu, Synchroextracting transform, *IEEE Trans. Ind. Electron.* **64**(10) (2017) 8042–8054.
31. T. Oberlin, S. Meignen and V. Perrier, The fourier-based synchrosqueezing transform, 2014 *IEEE Int. Conf. Acoustics, Speech and Signal Processing* (*ICASSP*) (2014), pp. 315–319.

32. K. He, X. Zhang, S. Ren and J. Sun, Deep residual learning for image recognition, 2016 *IEEE Conf. Computer Vision and Pattern Recognition* (*CVPR*) (2016), pp. 770–778.
33. J. Hammond and P. White, The analysis of non-stationary signals using time-frequency methods, *J. Sound Vib.* **190**(3) (1996) 419–447.
34. P. Jadhav, G. Rajguru, D. Datta and S. Mukhopadhyay, Automatic sleep stage classification using time–frequency images of cwt and transfer learning using convolution neural network, *Biocybern. Biomed. Eng.* **40**(1) (2020) 494–504.
35. B. Zhang, W. Wang, Y. Xiao, S. Xiao, S. Chen, S. Chen, G. Xu and W. Che, Cross-subject seizure detection in EEGs using deep transfer learning, *Comput. Math. Methods Med.* **2020** (2020).
36. Y. Cui and D. Wu, EEG-based driver drowsiness estimation using convolutional neural networks, in *Neural Information Processing*, eds. D. Liu, S. Xie, Y. Li, D. Zhao and E.-S. M. El-Alfy (2017), pp. 822–832.
37. Public Data set: CHB-MIT EEG dataset, Available at https://physionet.org/content/chbmit/1.0.0/ Accessed on 16 May 2020.
38. M. A. Ozdemir, M. Degirmenci, E. Izci and A. Akan, Eeg-based emotion recognition with deep convolutional neural networks, *Biomed. Eng. Biomed. Tech.* **66**(1) (2021) 43–57.
39. D. Gholamiangonabadi, N. Kiselov and K. Grolinger, Deep neural networks for human activity recognition with wearable sensors: Leave-one-subject-out cross-validation for model selection, *IEEE Access* **8** (2020) 133982–133994.
40. 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.* **99** (2018) 24–37.
41. R. Zhu, Y. Guo and J.-H. Xue, Adjusting the imbalance ratio by the dimensionality of imbalanced data, *Pattern Recogn. Lett.* **133** (2020) 217–223.
42. Y. Yuan, G. Xun, K. Jia and A. Zhang, A multi-context learning approach for EEG epileptic seizure detection, *BMC Syst. Biol.* **12**(6) (2018) 47–57.
43. W. Liang, H. Pei, Q. Cai and Y. Wang, Scalp EEG epileptogenic zone recognition and localization based on long-term recurrent convolutional network, *Neurocomputing* **396** (2020) 569–576.
44. M. S. Hossain, S. U. Amin, M. Alsulaiman and G. Muhammad, Applying deep learning for epilepsy seizure detection and brain mapping visualization, *ACM Trans. Multimedia Comput., Commun., Appl.* **15**(1s) (2019) 1–17.
45. Z. Wei, J. Zou, J. Zhang and J. Xu, Automatic epileptic EEG detection using convolutional neural network with improvements in time-domain, *Biomed. Signal Process. Control* **53** (2019) 101551.
46. M. Zhou, C. Tian, R. Cao, B. Wang, Y. Niu, T. Hu, H. Guo and J. Xiang, Epileptic seizure detection based on EEG signals and cnn, *Front. Neuroinformatics* **12** (2018) 95.
47. C.-L. Liu, B. Xiao, W.-H. Hsaio and V. S. Tseng, Epileptic seizure prediction with multi-view convolutional neural networks, *IEEE Access* **7** (2019) 170352–170361.
48. M. Shahbazi and H. Aghajan, A generalizable model for seizure prediction based on deep learning using cnn-lstm architecture, 2018 *IEEE Global Conf. Signal and Information Processing* (*GlobalSIP*) (2018), pp. 469–473.
49. O. K. Cura, M. A. Ozdemir and A. Akan, Epileptic EEG classification using synchrosqueezing transform with machine and deep learning techniques, 28*th European Signal Processing Conf.* (*EUSIPCO*) (IEEE, 2021), pp. 1210–1214.