Author(s): Wang, Xiaoshuang; Ristaniemi, Tapani; Cong, Fengyu

Title: One and Two Dimensional Convolutional Neural Networks for Seizure Detection Using EEG Signals

Year: 2020

Version: Accepted version (Final draft)

Please cite the original version:

Wang, X., Ristaniemi, T., & Cong, F. (2020). One and Two Dimensional Convolutional Neural Networks for Seizure Detection Using EEG Signals. In EUSIPCO 2020 : 28th European Signal Processing Conference (pp. 1387-1391). IEEE. European Signal Processing Conference. https://doi.org/10.23919/Eusipco47968.2020.9287640

One and Two Dimensional Convolutional Neural Networks for Seizure Detection Using EEG Signals

Xiaoshuang Wang[*†], Tapani Ristaniemi[†], Fengyu Cong[*†]
[*]*School of Biomedical Engineering, Dalian University of Technology, Dalian, China*
[†]*Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland*
xs.wang@foxmail.com, tapani.e.ristaniemi@jyu.fi, cong@dlut.edu.cn

***Abstract*—Deep learning for the automated detection of epileptic seizures has received much attention during recent years. In this work, one dimensional convolutional neural network (1D-CNN) and two dimensional convolutional neural network (2D-CNN) are simultaneously used on electroencephalogram (EEG) data for seizure detection. Firstly, using sliding windows without overlap on raw EEG to obtain the definite one-dimension time EEG segments (1D-T), and continuous wavelet transform (CWT) for 1D-T signals to obtain the two-dimension time-frequency representations (2D-TF). Then, 1D-CNN and 2D-CNN model architectures are used on 1D-T and 2D-TF signals for automatic classification, respectively. Finally, the classification results from 1D-CNN and 2D-CNN are showed. In the two-classification and three-classification problems of seizure detection, the highest accuracy can reach 99.92% and 99.55%, respectively. It shows that the proposed method for a benchmark clinical dataset can achieve good performance in terms of seizure detection.**

***Index Terms*—Electroencephalogram (EEG), seizure detection, convolutional neural networks (CNN), deep learning, time-frequency representation**

I. Introduction

Epilepsy is a chronic noncommunicable disease of the brain, which affects more than 50 million people worldwide. Clinically intractable epilepsy is commonly associated with the risk of fainting, injury, and death [1]. Electroencephalogram (EEG) is a significant tool that has been widely used for the diagnosis of epilepsy [2]. However, since the interpretation of EEG signals by visual assessment is labor- and time-consuming, the related research for EEG-based automatic seizure detection is very active. Automated and accurate identification of epileptic seizures based on EEG signals can improve work efficiency and patient quality of life.

The data mining techniques with feature-based engineering have been widely researched for the automated detection of epileptic seizures. Most of them use hand-wrought features extracted mainly from time domain [3], [4], time-frequency domain [5], [6], nonlinear dynamics [7], [8], and sometime in a combination of several domains [9] for seizures classification. However, these feature-based methods have several main challenges. First, EEG is non-stationary signal and can be susceptible to artifacts such as power-line interference, electrooculogram (EOG), electromyography (EMG) and white environment noise. All these noise sources can change the authenticity of features and hence seriously affect the performance of seizure detection systems. Second, feature extraction and selection has always been a time-consuming engineering. This is because the EEG data need to be processed, or further selected, to obtain desired features for classification.

Deep learning has proved its ability in image and audio recognition tasks [10], [11]. For solving the limitations mentioned above, convolutional neural network is used for the automated detection of seizures, and it is a machine learning technology based on representation learning. The system automatically learns and discovers the features needed for classification by processing multi-level input data [11]. In this work, one and two dimensional convolutional neural networks (1D-CNN and 2D-CNN) are used for seizure detection. Firstly, through preprocessing and continuous wavelet transform (CWT), we obtain the definite one-dimension time segments (1D-T) and two-dimension time-frequency representations (2D-TF), respectively. Then, two models are used on 1D-T and 2D-TF signals for classification, respectively. Finally, the classification results from two models are given. Two-classification and three-classification problems are discussed with the proposed method for seizure classification.

II. Data

The opening EEG data (http://epilepsy.uni-freiburg.de/database) collected by Andrzejak et al. [12] are used in this research. After removing EEG contaminated by artifacts (eye movements or muscle activity) through visual inspection, five sets (denoted A-E) of EEG data were selected. Each set contained 100 single-channel EEG segments of 23.6-sec duration. Sets A and B consisted of scalp EEG of five healthy volunteers with eyes open (A) and closed (B), respectively. The EEG signals in sets C and D were recorded from five patients during the seizure free intervals. Set C contained the EEG signals measured in the hippocampus, while the EEG signals of set D were measured in the epileptogenic zone. The EEG signals that are recorded during seizures in the epileptogenic zone were collected in set E. The sampling frequency is 173.61Hz with using 12bit resolution. A bandpass filter between 0.53 and 40Hz was used in the processing of recording. More details can be found in [12]. Fig. 1 shows the exemplary EEG signals of five sets A-E.

This work was supported by the National Natural Science Foundation of China (Grant No. 91748105 & 81471742), the Fundamental Research Funds for the Central Universities [DUT2019] in Dalian University of Technology in China and the scholarships from China Scholarship Council (No. 201806060166).

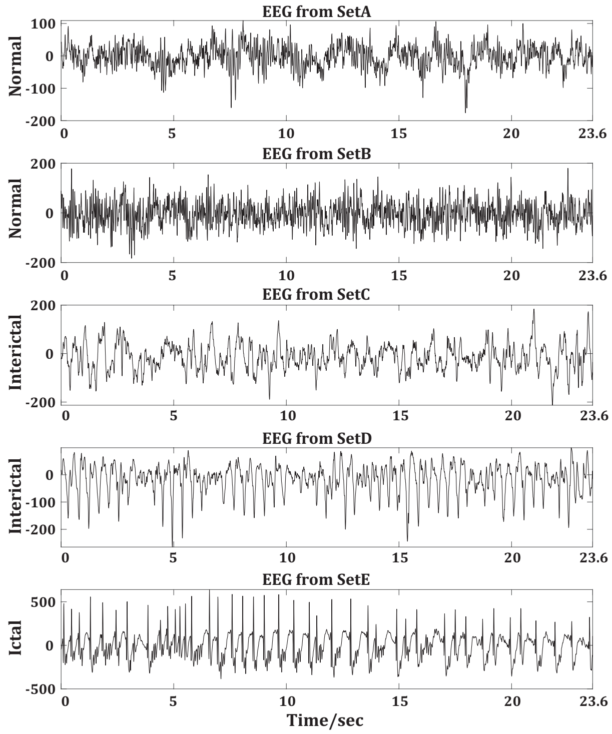

Fig. 1. Examples of EEG signals from each of the five sets of Bonn University EEG database.

III. Methodology

A. Preprocessing

Firstly, the 100 single-channel EEG segments from each set are integrated into one segment with length of 2360-sec duration. For obtaining different sample sizes, we then use two different sliding time windows (1-sec and 2-sec) without overlap to cut out the integrated EEG segments for attaining definite one-dimension time EEG fragments (1D-T) with lengths of 1-sec and 2-sec, respectively. In total, five integrated EEG segments A-E can obtain 11840 (1-sec) and 5920 (2-sec) samples, respectively.

B. Time-frequency representation

The changes of EEG signal are usually reflected in amplitude and frequency. Thus, time-frequency analysis is often used in abnormal EEG signals for seizure detection [5], [6]. Waveform transform is a commonly used time-frequency analysis method. For a EEG fragment $x(t)$, its time-frequency representation can be generated by using continuous wavelet transform (CWT), as follows:

$$\begin{aligned} \mathrm{TFR_x} &= |CWT_x(a,\tau)|^2 \\ &= |\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} x(t)\,\psi{*}(\frac{t-\tau}{a})dt|^2 \end{aligned} \tag{1}$$

where ψ and $*$ are the mother wavelet and function of complex conjugate, respectively. Parameters a and τ denote the oscillatory frequency and shifting position of the wavelet, respectively.

After attaining 1D-T signals by preprocessing, we then use CWT on them to generate the scalograms in time-frequency representation. In this paper, Morlet is used as mother wavelet to generate EEG time-frequency representations [6]. The time-frequency examples of EEG fragments from sets A-E are showed in Fig. 2.

C. Convolutional neural networks (CNN)

Fig. 3 shows the whole process of the proposed seizure detection system. Firstly, by preprocessing and time-frequency transform, we obtain the 1D-T and 2D-TF signals, respectively. Then, 1D-CNN is used for classifying 1D-T signals, while 2D-CNN is used for classifying the 2D-TF signals. Finally, the results from two CNNs are showed.

A CNN generally contains three types of layers: convolutional layer, pooling layer and fully connected layer. For a convolutional layer, it has a number of convolution kernels (filters) which perform convolution calculations on the input signals. Filtering results from convolution kernels are then nonlinearized by activation functions, such as rectified linear activation unit (ReLU) or Sigmoid functions. The output of a convolutional layer is usually known as the feature maps. Pooling layer is also called the down-sampling layer. Max-pooling operation is currently often used on the output from convolutional layer, it can preserve the most significant values from feature maps and improve the distortion tolerance of the model. After the operation of convolutional layers and pooling layers, the output is entered into the full connected layer. The results of classification are then achieved.

In this work, we use two different models (1D-CNN and 2D-CNN) and compare the results for exploring higher performance of seizure classification. About 1D-CNN, it contains three convolution blocks. Each convolution block consists of a convolution layer with ReLU as activation function, a batch normalization layer and a max-pooling layer. The three convolution layers have 16, 32 and 64 kernels, respectively, and kernels of the three convolution layers with same size of 3×1 and same stride of 1. Each batch normalization layer follows each convolution layer. Following each batch normalization layer is the max-pooling layer, with pooling size of 2. After the three convolution blocks, there are three fully connected layers. The first and second fully connected layers have 128 and 60 neurons with ReLU activation functions, respectively. The final layer is also known as output layer, and with 2 output neurons (two classification problem) or 3 output neurons (three classification problem). The Softmax activation function is used in the final layer. As showed in Fig. 3(a).

For 2D-CNN, it also contains three convolution blocks. Each convolution block also consists of a convolution layer with ReLU as activation function, a batch normalization layer and a max-pooling layer. The first convolution layer has 16 kernels with size of 3×3 and stride of 1×2. The second and third convolution layers have 32 and 64 kernels, respectively, and kernels of the two convolution layers with same size of 3×3 and same stride of 1×1. Each batch normalization layer follows each convolution layer. Following each batch

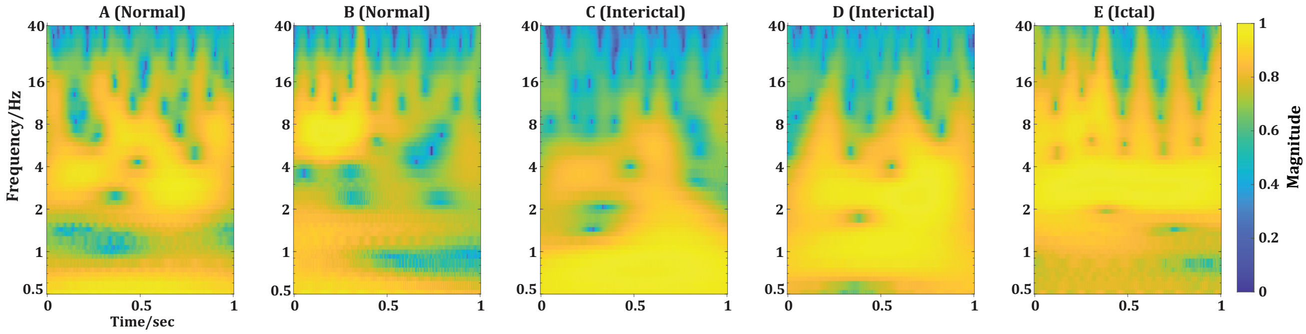

Fig. 2. Time-frequency representations of 1-sec EEG fragments from each of the five sets with using continuous wavelet transform.

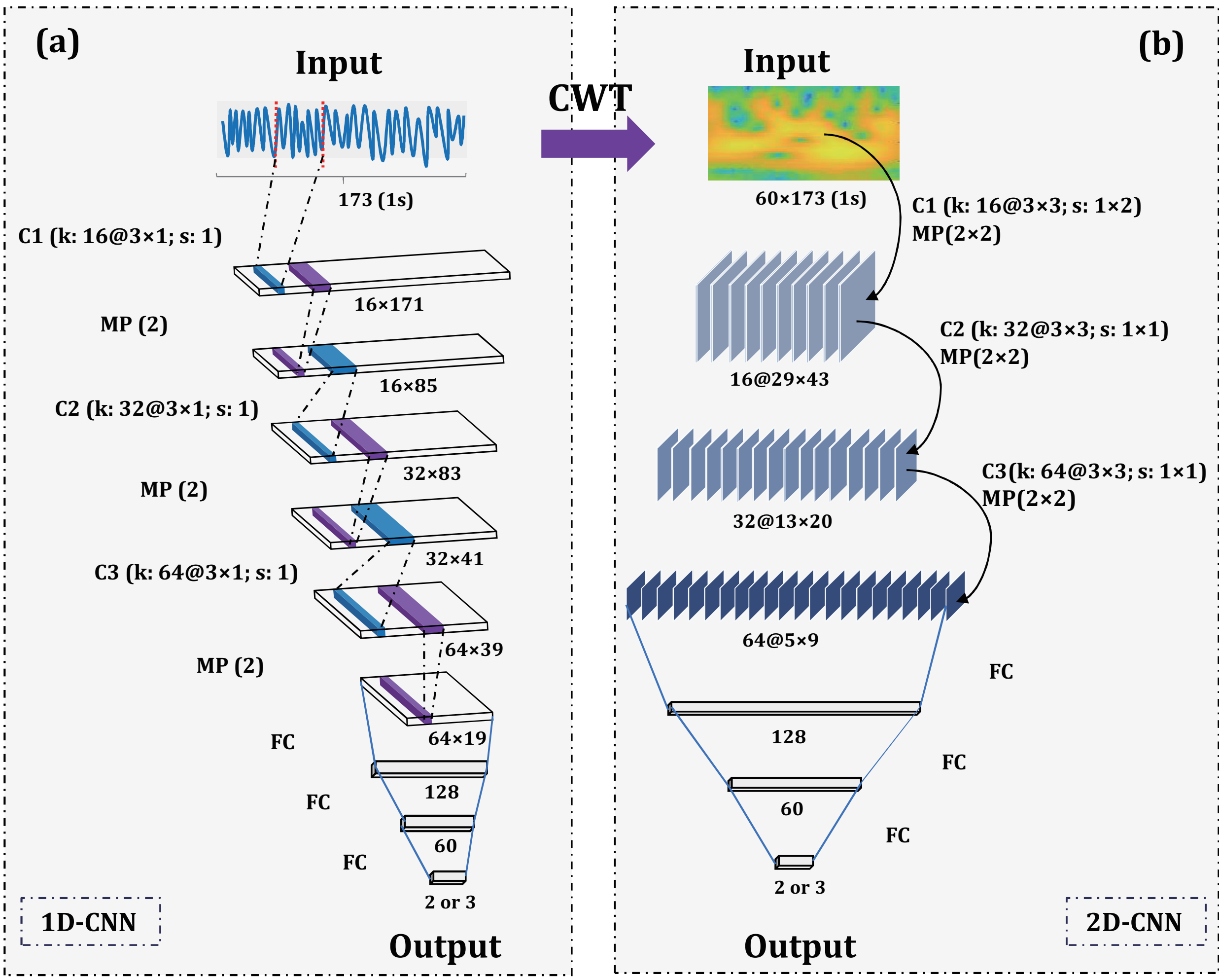

Fig. 3. Schematic diagram of the overall seizure detection approach with the preprocessing of 1-sec sliding time window. For simplicity, the batch normalization layers between the convolutional layers and the max-pooling layers are not shown. There are three convolution layers, named C1, C2, and C3. MP means max-pooling layer, and FC means full connected layer. The convolution kernel and stride are expressed as k and s, respectively.

normalization layer is the max-pooling layer, with pooling size of 2×2. Following the three convolution blocks are also three fully connected layers. The first and second fully connected layers also have 128 and 60 neurons with ReLU activation functions, respectively. The third layer has 2 output neurons or 3 output neurons with the Softmax activation function. As showed in Fig. 3(b).

D. *Training and testing of CNN models*

The ten-fold cross-validation method is used. All samples are first randomly divided into ten equal parts. Then, nine parts out of ten are used to train the CNN, while the remaining one is used to test the performance of trained CNN. This strategy is repeated ten times by shifting the test and training dataset.

TABLE I

THE CONFUSION MATRIX OF A VS. E ACROSS ALL TEN-FOLDS

	Normal (A)	Ictal (E)				Normal (A)	Ictal (E)			
	Predicted (1D-CNN, 1s)		**ACC (%)**	**SEN (%)**	**SPE (%)**	**Predicted (2D-CNN, 1s)**		**ACC (%)**	**SEN (%)**	**SPE (%)**
Original	2359	9	99.75	99.87	99.62	2368	0	**99.92**	**99.83**	**100**
	3	2365				4	2364			
	Predicted (1D-CNN, 2s)					**Predicted (2D-CNN, 2s)**				
	1184	0	**99.92**	**99.83**	**100**	1183	1	99.87	99.83	99.92
	2	1182				2	1182			

TABLE II

THE CONFUSION MATRIX OF C VS. E ACROSS ALL TEN-FOLDS

	Interictal (C)	Ictal (E)				Interictal (C)	Ictal (E)			
	Predicted (1D-CNN, 1s)		**ACC (%)**	**SEN (%)**	**SPE (%)**	**Predicted (2D-CNN, 1s)**		**ACC (%)**	**SEN (%)**	**SPE (%)**
Original	2343	25	99.01	99.07	98.94	2352	16	99.16	98.99	99.32
	22	2346				24	2344			
	Predicted (1D-CNN, 2s)					**Predicted (2D-CNN, 2s)**				
	1179	5	98.94	98.31	99.58	1177	7	**99.28**	**99.16**	**99.41**
	20	1164				10	1174			

TABLE III

THE CONFUSION MATRIX OF A VS. C ACROSS ALL TEN-FOLDS

	Normal (A)	Interictal (C)				Normal (A)	Interctal (C)			
	Predicted (1D-CNN, 1s)		**ACC (%)**	**SEN (%)**	**SPE (%)**	**Predicted (2D-CNN, 1s)**		**ACC (%)**	**SEN (%)**	**SPE (%)**
Original	2288	80	95.50	94.38	96.62	2243	125	93.69	92.65	94.72
	133	2235				174	2194			
	Predicted (1D-CNN, 2s)					**Predicted (2D-CNN, 2s)**				
	1161	23	**97.47**	**96.88**	**98.06**	1151	33	96.66	96.11	97.21
	37	1147				46	1138			

TABLE IV

THE CONFUSION MATRIX OF A VS. C VS. E ACROSS ALL TEN-FOLDS

	Normal (A)	Interictal (C)	Ictal (E)				Normal (A)	Interictal (C)	Ictal (E)			
	Predicted (1D-CNN, 1s)			**ACC (%)**	**SEN (%)**	**SPE (%)**	**Predicted (1D-CNN, 1s)**			**ACC (%)**	**SEN (%)**	**SPE (%)**
Original	2292	76	0	96.26	96.79	95.99	2209	157	2	95.03	93.29	95.90
	187	2171	10	95.99	91.68	98.14	191	2161	16	94.69	91.26	96.41
	3	12	2353	99.65	99.37	99.79	3	13	2352	99.52	99.32	99.62
	Predicted (1D-CNN, 2s)						**Predicted (2D-CNN, 2s)**					
	1146	36	2	96.96	96.79	97.04	1148	36	0	**96.93**	**96.96**	**96.92**
	70	1106	8	96.51	93.41	98.06	71	1107	6	**96.59**	**93.50**	**98.14**
	0	10	1174	99.44	99.16	99.58	2	8	1174	**99.55**	**99.16**	**99.75**

TABLE V

THE CONFUSION MATRIX OF AB VS. CD VS. E ACROSS ALL TEN-FOLDS

	Normal (AB)	Interictal (CD)	Ictal (E)				Normal (AB)	Interictal (CD)	Ictal (E)			
	Predicted (1D-CNN, 1s)			**ACC (%)**	**SEN (%)**	**SPE (%)**	**Predicted (1D-CNN, 1s)**			**ACC (%)**	**SEN (%)**	**SPE (%)**
Original	4550	183	3	96.12	96.07	96.16	4437	290	9	94.85	93.69	95.62
	261	4446	29	95.18	93.88	96.04	299	4356	81	93.86	91.98	95.12
	12	98	2258	98.80	95.35	99.66	12	57	2299	98.66	97.09	99.05
	Predicted (1D-CNN, 2s)						**Predicted (2D-CNN, 2s)**					
	2309	59	0	**97.33**	**97.51**	**97.21**	2283	83	2	96.96	96.41	97.33
	93	2266	9	**96.94**	**95.69**	**97.78**	90	2260	18	96.28	95.44	96.85
	6	20	1158	**99.41**	**97.80**	**99.81**	5	29	1150	99.09	97.13	99.58

IV. RESULTS

In this section, the results from two CNN models are showed, and two-class classification problem and three-class classification problem are discussed. For two-class classification problem, there are three cases discussed, namely A (normal) vs. E (ictal), C (interictal) vs. E (ictal), A (normal) vs. C (interictal). For three-class classification problem, two cases are discussed, namely A (normal) vs. C (interictal) vs. E (ictal), AB (normal) vs. CD (interictal) vs. E (ictal). The confusion matrix across all ten-folds in this paper is showed, and the accuracy (ACC), sensitivity (SEN), and specificity

(SPE) values are calculated.

A. Two-classification problem

Table I reports the classification of A and E. As showed in Table I, model 2D-CNN with preprocessing of 1-sec sliding time window and model 1D-CNN with 2-sec sliding time window have the same top results that the accuracy, sensitivity and specificity are 99.92%, 99.83% and 100%, respectively. For C and E, Table II shows that model 2D-CNN with preprocessing of 1-sec sliding time window has the best result with the accuracy of 99.28%, sensitivity of 99.16% and specificity of 99.41%. Table III shows the top result from model 1D-CNN with preprocessing of 1-sec sliding time window, it has the accuracy of 97.47%, sensitivity of 96.88% and specificity of 98.06%, respectively. From the results, we can see that the accuracy of A vs. E, C vs. E are all greater than 99%, while A vs. C only has more than 97% accuracy.

B. Three-classification problem

We also study the performance of the proposed method in classifying three distinct classes of EEG activities: normal, interictal, and ictal. Table IV gives the results of A vs. C vs. E. As it shows, the best overall classification result from model 2D-CNN with preprocessing of 2-sec sliding time window. It is observed that a high percentage of 96.96% of normal EEG signals are correctly classified as normal EEG signals with 3.04% of the EEG signals wrongly classified as interictal (3.04%) and ictal (0%) classes. For interictal EEG signals, only 93.50% of them are correctly classified as the interictal EEG signals, and a small percentage of 6.00% and 0.50% of them are wrongly classified as normal and ictal, respectively. Similarly, 99.16% of the ictal EEG signals are correctly classified as ictal class with 0.84% wrongly classified as normal (0.17%) and preictal (0.67%) classes.

The classification results of AB vs. CD vs. E are reported in Table V. Unlike the three classifications mentioned above, the normal (AB) and interictal (CD) groups have twice as many samples as the ictal (E) group. Model 1D-CNN with using 2-sec sliding time window for preprocessing has the best overall performance. For the normal EEG signals, a high percentage of 97.51% of them are accurately classified as the normal class, and a small percentage of 2.49% and 0% of them wrongly classified as the interictal and ictal classes, respectively. It is also observed that 95.69% of the interictal EEG signals are accurately classified as the interictal class with 4.31% of them are wrongly classified as normal (3.93%) and ictal (0.38%). Similarly, 97.80% of the ictal EEG signals are accurately divided into the ictal class with 2.20% wrongly classified as the normal (0.51%) and interictal (1.69%) classes.

V. Discussion and Conclusion

In this study, we discussed two-classification and three-classification problems in seizure detection. For exploring performance of classification, two models (1D-CNN and 2D-CNN) were used as classifiers. About two-classification problem, three cases, namely A (normal) vs. E (ictal), C (interictal) vs. E (ictal), A (normal) vs. C (interictal), were discussed. About three-classification problem, two cases, namely A (normal) vs. C (interictal) vs. E (ictal), AB (normal) vs. CD (interictal) vs. E (ictal), were discussed. For C vs. E and A vs. C vs. E, the performance of 2D-CNN was better than that of 1D-CNN, while for A vs. C and AB vs. CD vs. E, the performance of 1D-CNN was better. For A vs. E, the two models had the same best performance of classification. As shown in the five tables, A vs. E had the highest accuracy of 99.92% in two-classification problem, while A vs. C vs. E had the best overall classification result in three-classification problem.

Through preprocessing with using two different sliding windows (1-sec and 2-sec) for changing sample sizes, the results showed that two models have the good generalization and robustness in working well with the benchmark clinical dataset. The application of CNNs also has minimum feature engineering. All of these give medical staff more opportunities to efficiently and accurately detect the seizures, and help patients to improve the quality of life.

References

[1] L. Kuhlmann, K. Lehnertz, M. P. Richardson, B. Schelter, and H. P. Zaveri, "Seizure prediction—ready for a new era," *Nature Reviews Neurology*, vol. 14, no. 10, pp. 618–630, 2018.

[2] C. E. Elger and C. Hoppe, "Diagnostic challenges in epilepsy: seizure under-reporting and seizure detection," *The Lancet Neurology*, vol. 17, no. 3, pp. 279–288, 2018.

[3] R. Meier, H. Dittrich, A. Schulze-Bonhage, and A. Aertsen, "Detecting epileptic seizures in long-term human eeg: a new approach to automatic online and real-time detection and classification of polymorphic seizure patterns," *Journal of clinical neurophysiology*, vol. 25, no. 3, pp. 119–131, 2008.

[4] G. R. Minasyan, J. B. Chatten, M. J. Chatten, and R. N. Harner, "Patient-specific early seizure detection from scalp eeg," *Journal of clinical neurophysiology: official publication of the American Electroencephalographic Society*, vol. 27, no. 3, p. 163, 2010.

[5] K. Fu, J. Qu, Y. Chai, and Y. Dong, "Classification of seizure based on the time-frequency image of eeg signals using hht and svm," *Biomedical Signal Processing and Control*, vol. 13, pp. 15–22, 2014.

[6] Y. Yuan, G. Xun, K. Jia, and A. Zhang, "A multi-context learning approach for eeg epileptic seizure detection," *BMC systems biology*, vol. 12, no. 6, pp. 47–57, 2018.

[7] L. Guo, D. Rivero, and A. Pazos, "Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks," *Journal of neuroscience methods*, vol. 193, no. 1, pp. 156–163, 2010.

[8] S.-H. Lee, J. S. Lim, J.-K. Kim, J. Yang, and Y. Lee, "Classification of normal and epileptic seizure eeg signals using wavelet transform, phase-space reconstruction, and euclidean distance," *Computer methods and programs in biomedicine*, vol. 116, no. 1, pp. 10–25, 2014.

[9] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection," *IEEE transactions on biomedical engineering*, vol. 54, no. 9, pp. 1545–1551, 2007.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[12] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Physical Review E*, vol. 64, no. 6, p. 061907, 2001.