EPILEPTIC SPIKE DETECTION BY RECURRENT NEURAL NETWORKS WITH SELF-ATTENTION MECHANISM

Kosuke Fukumori † *Noboru Yoshida* ‡ *Hidenori Sugano* ‡‡ *Madoka Nakajima* ‡‡ *Toshihisa Tanaka* †§¶

† Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology
‡ Department of Pediatrics, Juntendo University Nerima Hospital
‡‡ Department of Neurosurgery, Juntendo University Hospital
§ Rhythm-Based Brain Information Processing Unit, RIKEN CBS
¶ Tensor Learning Unit, RIKEN AIP

ABSTRACT

Automated identification of epileptiform discharges in electroencephalograms (EEG) for the diagnosis of epilepsy can mitigate the burden of manual searches. Recent effective methods based on machine learning–based classification have used detection of candidate waveforms with signal processing and pattern matching as preprocessing, and this method can determine the overall performance. This paper thus considers a scenario where candidates are not detected; that is, we propose a recurrent neural network (RNN)–based self-attention model that can be fitted from the EEG segments generated without spike candidates being detected. In comparison with the state-of-the-art machine learning models that can be applied to EEG classification (LightGBM and EEGNet), the proposed model achieved higher performance (average accuracy: 90.2%). This result strongly suggests that the self-attention mechanism is suitable to automated identification of the epileptiform discharge in the EEG.

Index Terms— epilepsy, spike detection, epileptiform discharge, neural networks, electroencephalogram (EEG)

1. INTRODUCTION

Epilepsy is a chronic brain disease that causes seizures of disorientation and convulsions associated with excessive electrical excitation of neurons in the brain. It is estimated that there are 50 million people with epilepsy in the world, and the chronic lack of epilepsy specialists (epileptologists) is a severe problem. For example, in Japan, there are only 700 epileptologists, although there are one million patients [1]. Besides, since the diagnosis of epilepsy requires specialized knowledge or skill, epileptologists spend much of their time in specialized diagnosis of individual patients. This has motivated the development of an automated diagnostic aid to support epileptologists.

This work was supported in part by JST CREST (JPMJCR1784).

One of the steps in the diagnosis process is the measurement and analysis of the electroencephalogram (EEG). This is because the symptoms of epilepsy are determined based on a vital bio-marker called epileptiform discharges (epileptic spikes), which are often observed in the EEG [2]. Since the process of discovering these epileptic spikes is very complex, several automated detection methods have been advanced to support this process [3–6].

One of the most effective methods to automatically detect epileptic spikes is machine learning, especially supervised learning. In recent studies, there has been a tendency to create machine learning datasets through the following steps [7–10].

1. Detect the temporal locations of epileptiform discharge candidates in the EEG.
2. Have an epileptologist attach supervised labels—epileptic or nonepileptic—to the candidate waveforms.
3. Generate small interval EEG segments including the candidate waveforms based on the detected candidates.
4. Train a machine learning model with the EEG segments and the annotated labels.

For example, in the study of [7], candidate waveforms were first extracted based on EEG local maxima. Then, an annotator labeled the EEG maxima as epileptic or non-epileptic, and the EEGs including the peak points—300 ms before the peak and 700 ms after the peak—were cropped as the segments. These pairs of segments and labels were used to fit a random forest model. Similarly, in the studies [8, 10] which employ neural network-based method, one-second EEG segments containing candidate spikes are detected using the threshold value of EEG amplitude. Also, in the study of [9], a neurologist manually identified and annotated epileptic spikes and non-epileptic spikes that resembled epileptic spikes in the EEG recordings. Based on these labels, 0.2-second segments containing epileptic or non-epileptic spikes were extracted.

Then, some machine learning models such as k-nearest neighbors (k-NN), naive Bayes (NB), decision tree (DT), or SVM were fitted using these segments.

However, a method that requires the candidate detection step may lead to the exclusion of epileptic spikes or the inclusion of non-epileptic waveforms such as environmental noise. Also, in such methods with fixed spike locations within the segments, since the perfomance of the fitted model is influenced by the window taken around the peak, the segmentation constraints should be faithfully adhered to even in the prediction of the EEG of other unknown patients. That is, the fitting of the machine learning models is highly dependent on the manner in which the EEGs containing epileptic spikes are segmented.

In this paper, we hypothesize that a machine learning model that embeds a temporal self-attention mechanism [11] could extract locations and ranges of particular interest within a segment to detect epileptic spikes while simplifying the rules for EEG segment creation. To this end, this paper proposes a recurrent neural network (RNN) model with a self-attention mechanism that learns many EEG segments, including epileptic spikes, where the peak location is not aligned in segments. The results showed that without candidate spike detection, the proposed model achieved a higher classification performance than conventional classifiers used in other studies.

2. RELATED WORKS

Many studies of epileptic spike detection have trained machine learning models with segments generated with respect to the peak point of the spike [7]. In the study of [7], common epileptiform discharge (in the form of candidate epileptic spikes) was detected based on a threshold value. Each candidate spike was annotated as either epileptic or not and extracted from the EEG as a one-second segment containing 300 ms before the considered peak point and 700 ms after. These segments were then classified using a random forest model. In other studies [8], [7], candidate epileptic spikes detected using a threshold in the EEG were extracted as one-second segments. These segments were classified using an SVM or a feed-forward neural network model. Also, in the study of [9], 56-point segments around the locations of spikes were extracted from the EEG recorded at 256 Hz. Hence, a 0.2-second EEG recording around the location of the peak was segmented for each spike. Then, several machine learning models such as k-NN, NB, DT, or SVM were used to classify these segments. In the study of [12] as well as [9], 0.2-second segments were extracted to classify them using k-NN, DT, or SVM.

In these cases, the locations of the peaks were predefined in the segments, and each machine learning model were trained on them. Furthermore, the performance of the training model depended on the window being taken around the peak, which had to be determined appropriately. Therefore, the misalignment of the peak locations had a critical effect on the prediction performance of the model.

3. METHOD

3.1. Dataset

This paper constructs a dataset based from the EEG with epileptic events annotated in [10]. Table 1 summarizes the EEG recordings and the corresponding epileptic events. As this table shown, EEG recordings were collected from 50 patients (23 males, 27 females) with either childhood epilepsy with centrotemporal spikes (CECTS) [13] or focal epilepsy [13] at the Department of Pediatrics, Juntendo University Nerima Hospital. The data were recorded with the international 10–20 methods using the Nihon Koden EEG-1200 system. The sampling frequency was 500 Hz. This dataset was recorded and analyzed under approval from the Juntendo University Hospital Ethics Committee and the Tokyo University of Agriculture and Technology Ethics Committee.

In the study of [10], time instances corresponding to peaks (extrema) in 50 EEG recordings were automatically detected by a peak detector, and then the thresholded extrema were manually annotated as either two events (epileptic spike or non-epileptic spike) by five experts. For each EEG recording, this process produced a list of events, denoted by $\mathcal{E} = \{(t_i, y_i)\}$, where t_i denotes the time of the ith extremum and y_i denotes the corresponding event.

The above event list led to a dataset of labelled segments of EEGs constructed as follows. First, as shown in Fig. 1, each EEG recording was cropped every one-second from the top of the recording. We picked up the segments, $\boldsymbol{x}_j$, that cover any elements in event list $\mathcal{E}$, to form the dataset $\mathcal{D} = \{(\boldsymbol{x}_j, c_j)\}$, where c_j denotes the label (either epileptic or non-epileptic) defined as the following. If segment $\boldsymbol{x}_j$ includes

1. a single event (t_i, y_i), then c_j is assigned the same label as y_i;
2. multiple events and those events are identical, then c_j is assigned to the same label as those events;
3. multiple events and those events are not identical (say, at least one epileptic event), then c_j is assigned to the epileptic label.

Last, z-score normalization was applied with mean value and standard deviation for each segment.

3.2. RNN-based model with self-attention mechanism

In this paper, we propose a model that can be fitted from EEG segments that are not temporally aligned with spikes. Fig. 2 illustrates the architecture of the proposed RNN-based model.

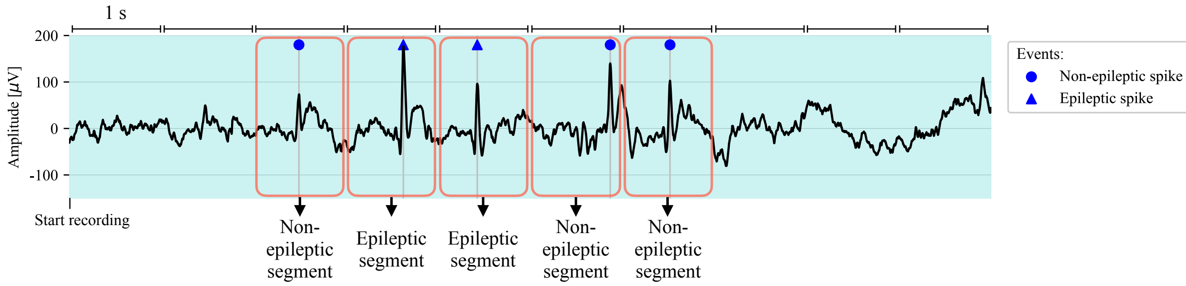


Fig. 1: Example of generating segments from the EEG recordings with supervised event labels.

Table 1: Dataset summary of 50 epileptic EEG [10]. This dataset was labeled by two neurosurgeons, two clinical technologists, and one pediatrician. The total number of labeled events (either epileptic spikes or non-epileptic discharges) is 31,486.

#Male patients	23
#Female patients	27
Age (Ave. ± STD.)	7.9 ± 2.0
Recording duration [min] (Ave. ± STD.)	27.5 ± 3.31
#Events of epileptic spikes	16,008
#Events of non-epileptic discharges	15,478
#Total ennotated events	31,486

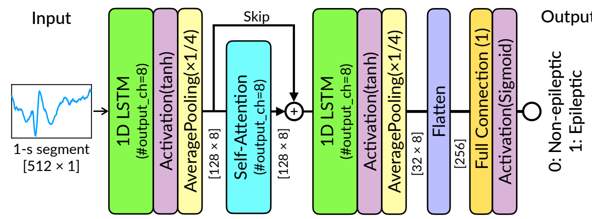


Fig. 2: Proposed 1D-RNN architecture with self-attention mechanism.

As this figure shows, the model mainly has two long short-term memory (LSTM) [14] layers and one self-attention layer. This paper adopts the dot-product attention layer [15] as this layer. That is, in this layer, three hidden features $\boldsymbol{Q}^{(i)} \in \mathbb{R}^{\tau \times d}$, $\boldsymbol{K}^{(i)} \in \mathbb{R}^{\tau \times d}$, and $\boldsymbol{V}^{(i)} \in \mathbb{R}^{\tau \times d}$, are calculated from the input features $\boldsymbol{X}^{(i)} \in \mathbb{R}^{\tau \times d}$ using the three weight matrices $\boldsymbol{W}_Q \in \mathbb{R}^{d \times d}$, $\boldsymbol{W}_K \in \mathbb{R}^{d \times d}$, and $\boldsymbol{W}_V \in \mathbb{R}^{d \times d}$ as

$$\boldsymbol{Q}^{(i)} = \boldsymbol{W}_Q \boldsymbol{X}^{(i)}, \tag{1}$$

$$\boldsymbol{K}^{(i)} = \boldsymbol{W}_K \boldsymbol{X}^{(i)}, \tag{2}$$

$$\boldsymbol{V}^{(i)} = \boldsymbol{W}_V \boldsymbol{X}^{(i)}, \tag{3}$$

where i, τ, and d are the index of EEG segments, the temporal length of the input feature, and the number of the feature channels, respectively. Then, using the softmax function and one weight matrix $\boldsymbol{W}_O \in \mathbb{R}^{d \times d}$, the output of the self-attention layer $\boldsymbol{Y}^{(i)} \in \mathbb{R}^{\tau \times d}$ is obtained as

$$\boldsymbol{Y}^{(i)} = \boldsymbol{W}_O \operatorname{softmax}\left(\boldsymbol{Q}^{(i)} (\boldsymbol{K}^{(i)})^{\top}\right) \boldsymbol{V}^{(i)}. \tag{4}$$

Therefore, it is expected that the first block of an LSTM and the average-pooling layers will output feature vectors every 8 ms, and the self-attention layer will search for relationships between these feature vectors. For the generation of initial weights of this model, the Xavier initializer [16] is used.

3.3. Experimental implementation

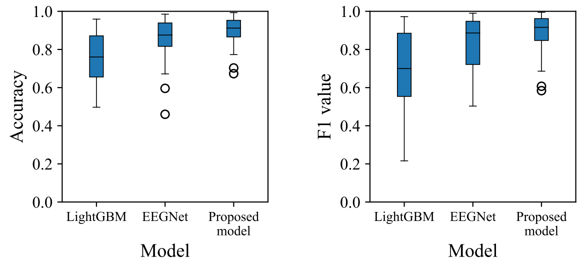
To verify the effectiveness of the proposed model, a numerical experiment was conducted using surface EEGs recorded from epileptic patients. Also, EEGNet [17] and LightGBM [18], which have been attractive for EEG analysis in recent years, are employed as comparison models. A one-second raw EEG segment is input to each model to classify the spike as epileptic or non-epileptic.

In this experiment, an intersubject validation was conducted with segments from 49 patients as train data and the remaining segments as test data in all combinations. A randomly selected 20% of the segments of the train data were used as validation data. EEGNet and the proposed model were fitted by the Adam optimizer [19], while overfitting was suppressed using early stopping [20] with the validation data. In the training of LightGBM, the number of estimators, which is its hyperparameter, was grid-searched from [5, 10, 20, 30, 50, 100, 300] using the validation data. To evaluate the models, the classification accuracy and the F1 value were adopted. When a dataset contains an imbalanced sample ratio—the sample ratio of a few patients is not balanced—the F1 value is a useful evaluation metric [21]. This can be calculated by the harmonic mean of precision and recall:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{5}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{6}$$

$$\text{F1} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{7}$$



(a) Accuracies (b) F1 values

Fig. 3: Graphical results of the 50 intersubject validations.

Table 2: Average accuracy and F1 value (average $\pm$ standard deviation) of the evaluation results in the intersubject validations. The best scores for each metric are marked in bold.

Model	Accuracy	F1 value
LightGBM	0.760 ± 0.127	0.702 ± 0.206
EEGNet	0.859 ± 0.102	0.821 ± 0.149
Proposed model	$\mathbf{0.902 \pm 0.068}$	$\mathbf{0.887 \pm 0.098}$

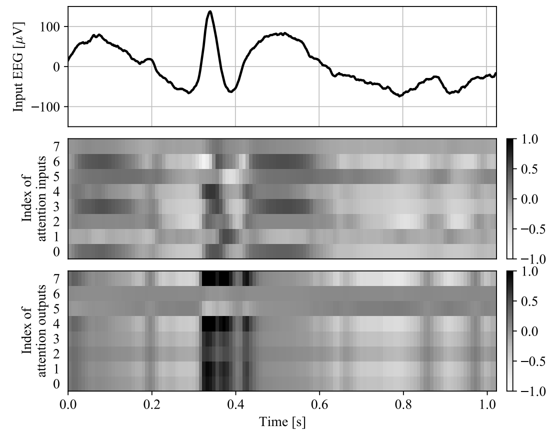
where TP, FP, FN, and TN are the numbers of true positive, false positive, false negative, and true negative, respectively.

4. EXPERIMENTAL RESULT

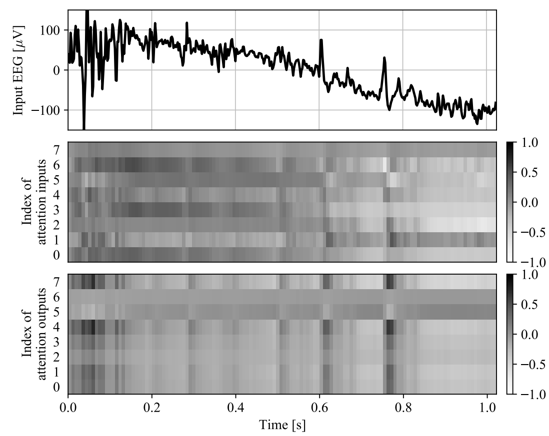
Fig. 3 graphically displays the accuracies and F1 values for each model in 50 intersubject validations, and Table 2 represents the average scores for each model. As shown in Table 2, the average accuracy and the average F1 value of the proposed model performed higher than the other models. Fig. 4 illustrates examples of the input features to the attention layer (namely, the 8-channel output of the recurrent layer) and their output features by the attention layer. Fig. 4a shows that feature extraction by recurrent layer tends to be affected by EEG amplitude. On the other hand, feature extraction by the attention layer responds more strongly to the temporal location of the epileptic spike. Also, the duration of the strong response seems to be approximately 100 ms. Similarly, Fig. 4b shows the features for non-epileptic segment input. This figure shows that even though the attentional layer responds to EEG maxima at around 600 ms and 750 ms, the responses are weaker and their duration is shorter than the response to epileptic spikes.

5. DISCUSSION

This paper has proposed a method to detect epileptic EEG segments that are not temporally aligned with spikes to eliminate the process of candidate spike detection. The proposed RNN-based model, combined with the self-attention mechanism, achieved higher detection performance than the conventional models (average accuracy: 90.2%). Therefore, although previous studies have generated segments based on the location of the spike candidates [7–9,12], it is fully detectable even if the spikes are not aligned within the segment. In other words, detection of candidate spikes was regarded as essential for machine learning spike detection; however, the results suggest that this detection may no longer be necessary. It is considered that the self-attention mechanism contributed to the improvement of accuracy by extracting the temporal location and range of interest accurately.



(a) Case of predicting an epileptic segment. The peak waveform at around 0.3 seconds is an epileptic spike.



(b) Case of predicting a non-epileptic segment.

Fig. 4: The top rows of each figure show the input EEG segments to the proposed model. The middle and bottom rows show the input features and output features of the self-attention layer when fed its segments (the size of the features is 128-length $\times$ eight-ch, but the temporal direction is aligned to the same length as the input segment).

Furthermore, the features extracted by the self-attention layer suggested that the required temporal length of the EEG is around 100 ms for epileptic spike detection using machine learning techniques. Also, the results of this paper, which aim to assist diagnosis, have the potential to indicate the waveform locations of particular interest to medical specialists.

6. REFERENCES

[1] T. J. E. Society, "Clinical Specialists Certified by JES," http://square.umin.ac.jp/jes, 2021, (Accessed in March, 2021).

[2] A. Parmeggiani, G. Barcia, A. Posar, E. Raimondi, M. Santucci, and M. C. Scaduto, "Epilepsy and EEG paroxysmal abnormalities in autism spectrum disorders," *Brain and Development*, vol. 32, pp. 783–789, July 2010.

[3] O. Ozdamar and T. Kalayci, "Detection of spikes with artificial neural networks using raw EEG," *Computers and Biomedical Research*, vol. 31, pp. 122–124, 1998.

[4] J. Algimantas, B. Gytis, B. Dmitrijus, and S. Ruta, "EEG analysis - automtic spike detection," *Nonlinear Analysis: Modelling and Control*, vol. 16, no. 4, pp. 375–386, 2011.

[5] A. R. Johansen, J. Jin, T. Maszczyk, J. Dauwels, S. S. Cash, and M. B. Westover, "Epileptiform spike detection via convolutional neural networks," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2016, pp. 754–758.

[6] L. Xuyen, L. Thanh, D. Viet, T. Long, N. Trung, and N. Thuan, "Deep learning for epileptic spike detection," *VNU Journal of Science: Comp. Science & Com. Eng*, vol. 33, no. 2, pp. 1–13, 2017.

[7] J. E. L. Douget, A. Fouad, M. M. Filali, J. Pyrzowski, and M. L. V. Quyen, "Surface and intracranial EEG spike detection based on discrete wavelet decomposition and random forest classification," *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 475–478, 2017.

[8] K. Fukumori, H. T. Thu Nguyen, N. Yoshida, and T. Tanaka, "Fully data-driven convolutional filters with deep learning models for epileptic spike detection," in *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 2772–2776.

[9] L. T. Thanh, N. T. A. Dao, N. V. Dung, N. L. Trung, and K. Abed-Meraim, "Multi-channel EEG epileptic spike detection by a new method of tensor decomposition," *Journal of Neural Engineering*, vol. 17, no. 1, pp. 016023, Jan. 2020.

[10] K. Fukumori, N. Yoshida, H. Sugano, M. Nakajima, and T. Tanaka, "Epileptic spike detection using neural networks with linear-phase convolutions," *IEEE Journal of Biomedical and Health Informatics*, 2021 (early access).

[11] R. M. Farsani and E. Pazouki, "A transformer self-attention model for time series forecasting," *Journal of Electrical and Computer Engineering Innovations (JECEI)*, vol. 9, no. 1, pp. 1–10, 2021.

[12] N. T. Dao, L. Thanh, N. Viet-Dung, N. Linh-Trung, and H. Le, "New feature selection method for multi-channel EEG epileptic spike detection system," *VNU Journal of Science: Computer Science and Communication Engineering*, vol. 35, Dec. 2019.

[13] W. D. Shields and O. C. Snead III, "Benign epilepsy with centrotemporal spikes," *Epilepsia*, vol. 50, no. s8, pp. 10–15, August 2009.

[14] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[15] T. Luong, H. Pham, and C. D. Manning, "Effective approaches to attention-based neural machine translation," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Sep. 2015, pp. 1412–1421, Association for Computational Linguistics.

[16] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, 2010, pp. 249–256.

[17] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces," *Journal of Neural Engineering*, vol. 15, no. 5, pp. 056013, Jul. 2018.

[18] P. Detti, G. Z. M. Lara, R. Bruni, M. Pranzo, F. Sarnari, and G. Vatti, "A patient-specific approach for short-term epileptic seizures prediction through the analysis of EEG synchronization," *IEEE Transactions on Biomedical Engineering*, vol. 66, pp. 1494–1504, Oct. 2019.

[19] M. Bernico, *Deep Learning Quick Reference : Useful hacks for training and optimizing deep neural networks with TensorFlow and Keras*, Packt Publishing Limited, 2018.

[20] D. Sarkar, R. Bali, and T. Ghosh, *Hands-On Transfer Learning with Python*, Packt Publishing Limited, 2018.

[21] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.