Satelight: self-attention-based model for epileptic spike detection from multi-electrode EEG

Kosuke Fukumori 1, Noboru Yoshida 1, Hidenori Sugano 3, Madoka Nakajima 3 and Toshihisa Tanaka 1,∗

1 Tokyo University of Agriculture and Technology, Koganei-shi, Tokyo, Japan
2 Juntendo University Nerima Hospital, Nerima-ku, Tokyo, Japan
3 Juntendo University School of Medicine, Bunkyo-ku, Tokyo, Japan
∗ Author to whom any correspondence should be addressed.
E-mail: tanakat@cc.tuat.ac.jp
Keywords: epilepsy, spike detection, electroencephalogram (EEG), self-attention

Abstract

Objective. Because of the lack of highly skilled experts, automated technologies that support electroencephalogram (EEG)-based epilepsy diagnosis are advancing. Deep convolutional neural network-based models have been used successfully for detecting epileptic spikes, one of the biomarkers, from EEG. However, a sizeable number of supervised EEG records are required for training. Approach. This study introduces the Satelight model, which uses the self-attention (SA) mechanism. The model was trained using a clinical EEG dataset labeled by five specialists, including 16 008 epileptic spikes and 15 478 artifacts from 50 children. The SA mechanism is expected to reduce the number of parameters and efficiently extract features from a small amount of EEG data. To validate the effectiveness, we compared various spike detection approaches with the clinical EEG data. Main results. The experimental results showed that the proposed method detected epileptic spikes more effectively than other models (accuracy = 0.876 and false positive rate = 0.133). Significance. The proposed model had only one-tenth the number of parameters as the other effective model, despite having such a high detection performance. Further exploration of the hidden parameters revealed that the model automatically attended to the EEG’s characteristic waveform locations of interest.

1. Introduction

Epilepsy is a chronic brain condition that results in seizures characterized by disorientation and convulsions associated with excessive electrical stimulation of brain neurons [1]. Epilepsy affects about 50 million people worldwide, and there is a severe shortage of specialists who treat epilepsy (epileptologists). For example, Japan only has 700 epileptologists, regardless of having one million patients suffering from epilepsy [2]. Furthermore, a lot of epileptologists’ work is necessary to care for each patient because the diagnosis of epilepsy requires specialized knowledge or skill. This has inspired the development of an automated diagnostic tool to aid epileptologists.

A crucial diagnostic step is the measurement and analysis of an electroencephalogram (EEG). This is because the epilepsy symptoms are determined based on a crucial biomarker called epileptiform discharges (epileptic spikes), which are often seen in a patient’s interictal EEG [3]. Because locating these epileptic spikes is difficult, numerous automated detection techniques have significantly improved over time to support this process [4, 5].

Particularly, neural network (NN)-based techniques have demonstrated high performance in spike detection tasks [6–10]. Epileptic spikes have been detected on a single-electrode EEG. Although the shape of the epileptic waves is similar among patients, each patient’s distribution of the waveforms is unique. To identify these individual differences in distribution among patients, specialists use several montages in epilepsy diagnosis, like bipolar and monopolar, to observe EEGs. This is because the specialists use the EEG of the surrounding or all electrodes to identify epileptic spikes. Therefore, to effectively use the detection methods from a single-electrode EEG, the appropriate montage must be...
determined manually, limiting its versatility. These facts have influenced studies on machine learning-based automatic spike detection. Since 2019, numerous related studies have used multi-electrode EEG to detect the location of epileptic spikes [11–15].

For detecting spikes using these multi-electrode EEG, convolutional NN (CNN)-based models, such as the model by Thomas et al’s [14] and SpikeNet [12], are some of the effective methods. Thomas et al [14] used CNN to detect epileptic spikes from multi-electrode EEG by calculating the probability of the appearance of epileptic spikes for individual electrodes and then obtaining the maximum of the output probabilities for all electrodes. SpikeNet [12] directly outputs a predicted value from a multi-electrode EEG segment. This was achieved using a combination of deep convolution layers in the spatial and temporal directions. SpikeNet efficiently extracts features using a top convolutional layer that merges the relationships between electrodes and 22 temporal convolutional layers. According to Jing et al [12], SpikeNet outperforms the detection performance of the commercially available software, Persyst13 [16], in spike detection.

However, training CNN-based models with many parameters necessitates a large labeled dataset, and gathering a significant amount of epileptic EEG is challenging [17]. Therefore, developing a lightweight model suitable for EEG analysis can contribute to epileptic spike detection from multi-electrode EEG. Here, the analysis of a multi-dimensional time series can consider spike detection from a multi-electrode EEG. According to Farsani and Pazouki [18], the number of parameters in time-series data analysis can be reduced using a mechanism called self-attention (SA). By learning the relationship between all sampled points in a segment, the SA mechanism can efficiently discover significant features. Many convolution and pooling layers are required to implement this with CNNs. In this study, we hypothesized that the trainable parameters in the model for epileptic spike detection from EEG would be reduced using the SA mechanism. To this end, we constructed a spike detection model for multi-electrode EEG called the self-attention-based lightweight model (Satelight), which is built with a minimal set of parameters. For the verification of this method, different spike detection approaches were compared using a dataset consisting of clinical EEGs recorded from 50 children with either childhood epilepsy with centro-temporal spikes (CECTS) [19] or focal epilepsy, with supervised labels indicating the location of epileptic spikes annotated by five experts.

2. Method

2.1. Dataset

Table 1 summarizes the dataset. Interictal EEG recordings were collected from 50 children (23 males and 27 females) with either CECTS or focal epilepsy [19] at the Department of Pediatrics, Juntendo University Nerima Hospital. The patients’ ages at the time of the examination ranged between three and twelve years. The data were recorded from 16 electrodes with the International 10–20 system using the Nihon Koden EEG-1200 system. The sampling frequency was 500 Hz. This dataset was recorded and analyzed with the approval from the Juntendo University Hospital Ethics Committee and the Tokyo University of Agriculture and Technology Ethics Committee.

Measurement EEG recordings were pre-processed as follows. First, using Scipy [20], a peak search method was used to find the EEG’s peak waveforms (minima and maxima) with. With a minimum distance of 100 points, this function extracts both the upward and downward peaks. Meaningless peaks caused by noise and other factors were removed using a threshold set at the 80th percentile of all peaks in one electrode. Second, for the peaks (namely, candidate spikes) that were actual epileptic spikes (spike or spike-and-wave), five experts annotated labels as epileptic. To produce accurate non-epileptic segments, the experts also labeled non-epileptic discharges. These expert-selected non-epileptic discharges, which were mistakenly detected as candidate spikes, included electromyograms, ambient noise, and alpha EEG. Finally, the EEG recordings were cropped every 1.024 s (512 sampling points) from the beginning of the recordings, independent of the temporal locations of the labeled events. As shown in figure 1, each segment that contained either epileptic or non-epileptic labeled EEG was extracted as a valid segment. According to figure 1, segments with one or more epileptic spikes are considered epileptic segments even if they have non-epileptic labels, as well. Thus, one segment had 16 electrodes and 512 samples (1.024 mathrm s). Z-score normalization was used with the mean value and standard deviation for each segment.

2.2. Separable convolution

A layer known as separable convolution, which independently convolves the temporal and spatial directions, should be inserted at the beginning of the network for the analysis of multi-electrode EEG with NNs. This involves frequency filtering for the
temporal convolution and electrode coupling for the spatial convolution (i.e. montage optimization). Given that it has a separate kernel for each dimension, this convolutional layer has much smaller parameters than a standard two-dimensional convolutional layer. The first model to use this layer for EEG analysis is EEGNet [21], which has been successful in the field of brain-computer interface. Huang et al [22] also reported the effectiveness of the EEGNet-based model in their study of EEG-based emotion classification tasks. Using an extended model of EEGNet, Shoji et al [23] successfully identified abnormal EEG durations indicative of patients with juvenile absence epilepsy. Furthermore, SpikeNet [12] uses a separable convolution to detect epileptic spikes. As these results show, the separable convolution in multi-electrode EEG analysis is effective, particularly in extracting features along the electrode direction.

2.3. Self-attention mechanism

Accurately detect epileptic spikes, even in randomly extracted EEG segments, without using candidate detectors. However, to train a deep learning model with many parameters [17], a significant amount of training data are needed. We focused on the SA mechanism to efficiently analyze the waveform of interest. The SA mechanism is expected to automatically skip redundant data in time-series signals [18]. Using the dot-product attention [24] is one way to construct SA. That is, the layer based on SA calculates three hidden features, \( Q^{(i)} \in \mathbb{R}^{r \times d} \), \( K^{(i)} \in \mathbb{R}^{r \times d} \), and \( V^{(i)} \in \mathbb{R}^{r \times d} \), from the input features \( X^{(i)} \in \mathbb{R}^{r \times d} \) using the three weight matrices, \( W_Q \in \mathbb{R}^{d \times d} \), \( W_K \in \mathbb{R}^{d \times d} \), and \( W_V \in \mathbb{R}^{d \times d} \), as follows:

\[
Q^{(i)} = X^{(i)} W_Q, \quad K^{(i)} = X^{(i)} W_K, \quad V^{(i)} = X^{(i)} W_V, \quad (1-3)
\]

As (4) shows, the softmax function of the matrix product between \( Q^{(i)} \) and \( K^{(i)} \) acts to represent the level of interest within the input feature itself. Further, it is expected that multiplying the softmax using the weighted input feature \( V^{(i)} \) extracts meaningful EEG locations for the prediction. Here, the SA layer can be constructed as a single layer using these calculations.

2.4. Satelight: a self-attention-based lightweight model

In this study, we propose a model that can be trained using multi-electrode EEG segments and labels that indicate whether the spikes are epileptic. Figure 2 illustrates the architecture of the proposed Satelight. As shown in figure 2, the model is factorized into the first convolution block and three SA blocks. The first convolution block combines temporal and spatial convolutions, whose effectiveness in lowering the number of learnable parameters has been reported in [12, 21]. A two-dimensional convolution layer with a kernel size of \((f_s/2, 1)\) implements the temporal convolution. Because the temporal kernel size is half that of the sampling frequency, it theoretically constitutes a filter with a frequency resolution of 2 Hz [21]. Spatial convolution is performed using a two-dimensional depthwise convolution layer. In this layer, 16 input matrices of size \((N, C)\) are independently convolved with \(C\)-sized kernels. Thus, it is expected that the desired electrodes will be selected. These convolutions feed the following SA blocks with the features that have been spatiotemporally filtered. Therefore, every 8 ms, the SA layer is expected to search for relationships between feature vectors. Finally, the fully connected layer with sigmoid activation is adopted as the output layer for binary classification. For training Satelight, cross-entropy, which
Figure 2. Architecture of the proposed Satelight, where \( N \), \( C \), and \( f_s \) are the length of the input segment, the number of EEG electrodes, and the sampling frequency, respectively. In this study, \( N = 512 \), \( C = 16 \), and \( f_s = 500 \).

\[
\text{Figure 2. Architecture of the proposed Satelight, where } N, C, \text{ and } f_s \text{ are the length of the input segment, the number of EEG electrodes, and the sampling frequency, respectively. In this study, } N = 512, C = 16, \text{ and } f_s = 500.\]

is the log-likelihood of the Bernoulli distribution, is employed as the loss function:

\[
J = -l(i) \log p(i) - (1 - l(i)) \log (1 - p(i)),
\]

where \( l(i) \in \{0, 1\} \) and \( p(i) \in [0, 1] \) denote the annotated label and the model’s prediction of the \( i \)th EEG segment, respectively.

2.5. Experimental implementation

To verify the effectiveness of the proposed Satelight, we conducted a numerical experiment using surface EEGs recorded from patients with epilepsy. As mentioned, this experiment is a binary classification of segments that contain or do not contain epileptic spikes. The following five NN-based models, one statistical property-based model, and one commercially available software are used as comparison models:

- **NN-based models:**
  - The proposed model (Satelight);
  - Thomas et al’s model \([14]\);
  - Two models proposed by Jing et al \([12]\) with the kernel size of all temporal convolutional layers set to:
    + three (SpikeNet3);
    + five (SpikeNet5);
  - Lawhern et al’s model (EEGNet) \([21]\);

- **Statistical property-based model:**
  - Janca et al’s model \([25]\);

- **Commercially available software:**
  - Persyst13 software \([16]\).

In this experiment, 50 leave-one-patient-out tests were performed using EEG segments from 49 patients as a training set and the remaining segments from one patient as a test set in all possible combinations. The Adam optimizer \([26]\) used the training set to train the models for comparison. Additionally, early stopping \([27]\) was applied using a portion of the training set (namely, the validation set) to suppress overfitting. The validation set consists of 20% randomly selected segments from the training data. Thus, 80% of the segments were used for training in the 49 EEG recordings. The Xavier initializer \([28]\) was used to generate the initial weights for this model. Note that the training process was not conducted on Persyst13 software.

To compare the classification performances, we used four evaluation criteria: the area under the curve (AUC) of the receiver operating characteristic (ROC), F1-value, true positive rate (TPR, sensitivity), and false-positive rate (FPR). F1-value, TPR, and FPR are calculated as follows:

\[
\text{TPR} = \frac{TP}{TP + FN},
\]
show heat maps of the first SA layer’s output (the size is [128, 32]), where the SA structure is illustrated in figure 2, and the upper EEG segments are input. As shown in figure 4, the epileptic spikes appeared between 0.1–0.2 s. The SA output also responded strongly in this range. Similarly, as shown in figure 5, the SA responded strongly to the epileptic spikes around 0.7 s. In this case, it did not respond to the artifacts seen around 0.3 s.

4. Discussion

4.1. Comparison of the number of parameters in the models

Table 3 and figure 6 show the number of parameters used in the numerical experiment for the compared models. EEGNet is a model with few parameters; however, as shown in table 2, its detection performance is not as high as that of SpikeNet and Satelight. The model by Thomas et al which had the highest number of parameters, produced a high FPR, as shown in table 2. This might be because, although it is a deep network, it does not extract features in the spatial direction. Compared with these two models, SpikeNet is more stable in detecting epileptic spikes (accuracy = 0.860 and F1 = 0.821 using SpikeNet5). A combination of deep convolution layers in both the spatial and temporal directions could have been achieved used to achieve this. Moreover, even though Satelight performed better at classification than SpikeNet5 (accuracy = 0.876 and F1 = 0.843), it can be constructed with one-tenth as many parameters. This implies that the deep convolutional layer may be redundant. Therefore, using the SA layer for epileptic spike detection is effective.

4.2. Analysis of self-attention layer behavior

The SA layer behavior analysis results, such as the strong response during epileptic spike and no response on artifacts, as shown in figures 4 and 5 may imply that SA operates independently of the EEG amplitude. Moreover, these heat maps demonstrate that the SA responses are slightly delayed (≈200 ms) from the onset of the spike. This delay appears to be caused by specialists’ focus on the positive phase reversal following the spike [32]. According to the findings, this experiment, which classifies epileptic
spikes and artifacts, does not depend on the EEG near the beginning of the event. Therefore, the behavior of the SA may change if the task involves identifying specific epileptic spike types, such as a spike–wave or sharp–wave.

4.3. Dataset collection
Generally, a dataset should include a variety of age, gender, and epilepsy types, as well as a large amount of data and accurate annotations. In this study, as many as 50 EEG recordings were collected for reliable
validation. Additionally, the $Z$-score normalization was applied as preprocessing step to the EEG segments to reduce amplitude differences among individuals. Note that spectralization as preprocessing step is also available [33], however, this paper did not adopt it owing to the increase in configurations to be adjusted—such as window size and window function—and the increases in input dimensions and learnable model parameters. We consider that the proposed model trained on this dataset has high versatility for typical epileptic spike detection. To further improve versatility, the correctness of annotations must be examined. In this study, all samples were annotated at the sole discretion of the annotators without consulting other annotators. That is, the annotation process may contain human errors. The SpikeNet detection result in our experiment (AUC = 0.92) was less accurate than the original result (AUC = 0.98) reported by Jing et al [12]. One possible reason is that only samples that six or more annotators classified as epileptic were used in their evaluation.

In this study, the age range of the dataset was limited to children. Although it has not been reported that the typical form of epileptic spikes varies with age and gender; however, it is essential to collect patient data from a variety of age groups for reliable validation. Additionally, to build a more practical model, other epileptiforms, such as quasi-periodic spikes and poly-spike, must be annotated. To solve these limitations, it is desirable to develop a framework for more efficient data collection and labeling.

**4.4. Segment extraction for dataset construction**

In many spike detection studies, after detecting candidate spikes and annotating them with labels for the dataset construction, the EEG signals are segmented out based on the labeled spikes. Candidate spikes are typically found using simple signal processing methods and expert selection directly from the EEG. Single-electrode EEG studies use the former method [8, 10, 34]. The latter method is used for both single and multi-electrode EEG studies [12–14]. Because the candidate spikes are not uniquely defined, there is currently no simple segment method for multi-electrode EEG extraction. For example, detecting peak spikes based on amplitude is challenging because spikes do not always occur at precisely the same time over multiple electrodes. Furthermore, depending on epilepsy symptoms, the number of electrodes...
Figure 5. Example of a hidden feature map of the first SA layer when predicting an epileptic segment with extreme artifacts at approximately 0.3 s. For better visualization, the 128 temporal points in the feature map are stretched as 1.024 s. The peak waveform at around 0.7 s is an epileptic spike.

Table 3. Number of parameters of the comparison models.

| Model                        | No. of parameters | Ratio with Satelight |
|------------------------------|-------------------|----------------------|
| Satelight (proposed model)   | 83,905            | ×1.00                |
| EEGNet [21]                  | 3061              | ×0.0365              |
| Thomas et al's model [14]    | 163,875,784       | ×195                 |
| SpikeNet3 [12]               | 518,658           | ×6.18                |
| SpikeNet5 [12]               | 848,386           | ×10.1                |

Figure 6. A visual comparison of the model parameters’ numbers. The proposed model (Satelight) is lighter than the other models except for EEGNet, even though it achieves the highest accuracy of all models.
in which spikes appear varies; thus, the rules for candidate detection are likely to be complex.

However, if segments are extracted based on the expert’s selected spikes, the validation data appear to depend on the expert’s decision even, and then a data set is constructed using these segments. Therefore, for sufficient validation, the expert’s decision must be removed from the factors in the validation phase. Following this justification, the dataset for this study was created using the following methodology:

(a) Split the EEG recordings as segments at fixed intervals from the beginning.
(b) Indicate on labels whether the segment contains epileptic spikes.

The method for extracting segments of the validation data using this procedure was superior to the method relying on expert selection. The disadvantage of this method is that the location of spikes varies for each segment; thus, a more effective method for segment extraction or identifying characteristic waveforms for multi-electrode EEG may be necessary. In this study, the SA mechanism, as described in section 3.2, rather than the segment extraction method, successfully mitigated this limitation.

4.5. Implementation for practical use

In this study and numerous others, datasets were built using pairs of EEG segments and corresponding labels, then divided into training and test sets for verification [8, 10, 12–14, 34]. This validation will be sufficient if spike detection can be considered a simple classification task. However, this validation method using a pre-constructed dataset is a limitation of these studies because the raw EEG recordings are not segmented. Therefore, ideally, all EEG recordings should be fully labeled by the experts without mistakes or omissions. Alternatively, detecting spikes from the entire EEG recordings and then evaluating these predictions posteriorly, rather than constructing the test set, is preferable. In other words, for practical use, a fully automated method for detecting spikes in raw EEG recordings, including EEG segmentation steps, and evaluating them appropriately is urgently required.

5. Conclusion

We propose a lightweight epileptic spike detection model that employs the SA mechanism. The number of parameters in this model is small compared with that of the other state-of-the-art deep NN-based models. Nevertheless, the model achieved high detection performance (accuracy $= 0.876$ and FPR $= 0.133$). Furthermore, an exploration of the hidden parameters showed that the model automatically paid attention to the characteristic waveform locations of interest in the EEG. This would significantly contribute to the development of an explainable NN.

Data availability statement

The data generated and/or analysed during the current study are not publicly available for legal/ethical reasons but are available from the corresponding author on reasonable request.

Acknowledgments

The authors thank Ms Meiko Sakurai and Ms Junko Hirota for their help with annotation as well as Dr Miao Yao and Mr Kazi Mahmudul Hassan for an insightful discussion on technical analysis. This work was supported by JST CREST Grant Number JPMJCR1784. We would like to thank Editage (www.editage.com) for English language editing.

ORCID iDs

Kosuke Fukumori https://orcid.org/0000-0003-0389-9165
Toshihisa Tanaka https://orcid.org/0000-0002-5056-9508

References

[1] WHO 2021 Epilepsy (available at: www.who.int/news-room/fact-sheets/detail/epilepsy) (Accessed June 2021)
[2] Society T J E 2021 Clinical specialists certified by JES (available at: square.umin.ac.jp/jes) (Accessed March 2021)
[3] Parmeggiani A, Barcia G, Posar A, Raimondi E, Santucci M and Scaduto M C 2010 Epilepsy and EEG paroxysmal abnormalities in autism spectrum disorders Brain Dev. 32 783–9
[4] Oudamar O and Kalayci T 1998 Detection of spikes with artificial neural networks using raw EEG Comput. Biomed. Res. 31 122–4
[5] Algimantas J, Gybitis B, Dmitrijus B and Ruta S 2011 EEG analysis—automatic spike detection Nonlinear Anal.: Modelling Control 16 375–86
[6] Johansen A R, Jin J, Maszczuk T, Dauwels J, Cash S S and Westover M B 2016 Epileptiform spike detection via convolutional neural networks 2016 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP) pp 754–8
[7] Xuyn L, Thanh L, Viet D, Long T, Trung N and Thuan N 2017 Deep learning for epileptic spike detection VNU J. Sci. Comput. Sci. Commun. Eng. 33 1–13
[8] Douget J E L, Fouad A, Filali M M, Pyrzowski J and Quyen M L V 2017 Surface and intracranial EEG spike detection based on discrete wavelet decomposition and random forest classification 39th Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC) pp 475–8
[9] Bagheri E, Jin J, Dauwels J, Cash S and Westover M B 2019 A fast machine learning approach to facilitate the detection of interictal epileptiform discharges in the scalp electroencephalogram J. Neurosci. Methods 326 108362
[10] Fukumori K, Thu Nguyen H T, Yoshida N and Tanaka T 2019 Fully data-driven convolutional filters with deep learning models for epileptic spike detection 2019 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP) pp 2772–6
[11] Dao N T, Thanh L, Viet-Dung N, Linh-Trung N and Le H 2019 New feature selection method for multi-channel EEG
epileptic spike detection system VNU J. Sci. Comput. Sci. Commun. Eng. 35 47–59
[12] Jing J et al 2020 Development of expert-level automated detection of epileptiform detection by a new method of tensor decomposition J. Neural Eng. 17 016023
[13] Thanh T, Dan N T A, Dung N V, Trung N L and Abed-Meraim K 2020 Multi-channel EEG epileptic spike detection from scalp electroencephalograms by convolutional neural networks Int. J. Neural Syst. 30 2050030
[14] Thomas J, Jin J, Thangavel P, Bagheri E, Yuvaraj R, Daauwels J, Rathakrishnan R, Halford J J, Cash S and Westover B 2020 Automated detection of interictal epileptiform discharges in scalp electroencephalograms using separable convolution neural network with bilinear interpolation IEEE Access 8 131636–46
[15] Wei B, Zhao X, Shi L, Xu L, Liu T and Zhang J 2021 A deep learning framework with multi-perspective fusion for interictal epileptiform discharges detection in scalp electroencephalogram J. Neural Eng. 18 046003
[16] Scheuer M I, Bagic A and Wilson S B 2017 Spike detection: inter-reader agreement and a statistical Turing test on a large data set Clin. Neurophysiol. 128 243–50
[17] Lourenço C, Tjepkema-Cloostermans M and Putten M 2021 Efficient use of clinical EEG data for deep learning in epilepsy Clin. Neurophysiol. 132 1234–40
[18] Farsani R M and Pazouki E 2021 A transformer self-attention model for time series forecasting J. Electr. Comput. Eng. Innov. 9 1–10
[19] Shields W D and Sneed O C I I 2009 Benign epilepsy with centrotemporal spikes Epilepsia 50 10–15
[20] Virtanen P et al (SciPy 1.0 Contributors) 2020 SciPy 1.0: fundamental algorithms for scientific computing in Python Nat. Methods 17 261–72
[21] Lawhern V J, Solon A J, Waytowich N R, Gordon S M, Hung C P and Lance B J 2018 EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces J. Neural Eng. 15 056013
[22] Huang W, Xue Y, Hu L and Lu L 2020 S-EEGNet: Electroencephalogram signal classification based on a separable convolution neural network with bilinear interpolation IEEE Access 8 131636–46
[23] Shoji T, Yoshida N and Tanaka T 2021 Automated detection of abnormalities from an EEG recording of epilepsy patients with a compact convolutional neural network Biomed. Signal Process. Control 70 103013
[24] Luong T, Pham H and Manning C D 2015 Effective approaches to attention-based neural machine translation Proc. 2015 Conf. on Empirical Methods in Natural Language Processing (Association for Computational Linguistics) pp 1412–21
[25] Janca R et al 2014 Detection of interictal epileptiform discharges using signal envelope distribution modelling: application to epileptic and non-epileptic intracranial recordings Brain Topogr. 28 172–83
[26] Bernico M 2018 Deep Learning Quick Reference: Useful Hacks for Training and Optimizing Deep Neural Networks With TensorFlow and Keras (Birmingham: Packt Publishing Limited)
[27] Sarkar D, Bali R and Ghosh T 2018 Hands-On Transfer Learning With Python (Birmingham: Packt Publishing Limited)
[28] Glorot X and Bengio Y 2010 Understanding the difficulty of training deep feedforward neural networks Proc. 13th Int. Conf. on Artificial Intelligence and Statistics pp 249–56
[29] Chollet F et al 2015 Keras (available at: https://keras.io)
[30] Pedregosa F et al 2011 Scikit-learn: machine learning in Python J. Mach. Learn. Res. 12 2825–30
[31] Siegel S and Castellan N 1988 Nonparametric Statistics for the Behavioral Sciences 2nd edn (New York: McGraw–Hill, Inc.)
[32] Libenson M 2010 Practical Approach to Electroencephalography (Philadelphia, PA: Saunders)
[33] Nejedly P et al 2019 Intracerebral EEG artifact identification using convolutional neural networks Neuroinformatics 17 225–34
[34] Fukumori K, Yoshida N, Sugano H, Nakajima M and Tanaka T 2022 Epileptic spike detection using neural networks with linear-phase convolutions IEEE J. Biomed. Health Inf. 26 1045–56