sensors

MDPI

Article

Automatic Seizure Detection Based on Stockwell Transform and Transformer

Xiangwen Zhong [1], Guoyang Liu [1], Xingchen Dong [1], Chuanyu Li [1], Haotian Li [1], Haozhou Cui [1] and Weidong Zhou [1,2,*]

[1] School of Integrated Circuits, Shandong University, Jinan 260100, China; 202212333@mail.sdu.edu.cn (X.Z.); gyliu@sdu.edu.cn (G.L.); 202132378@mail.sdu.edu.cn (X.D.); 202232355@mail.sdu.edu.cn (C.L.); 202212327@mail.sdu.edu.cn (H.L.); 202232343@mail.sdu.edu.cn (H.C.)

[2] Shenzhen Institute, Shandong University, Shenzhen 518057, China

* Correspondence: wdzhou@sdu.edu.cn

Abstract: Epilepsy is a chronic neurological disease associated with abnormal neuronal activity in the brain. Seizure detection algorithms are essential in reducing the workload of medical staff reviewing electroencephalogram (EEG) records. In this work, we propose a novel automatic epileptic EEG detection method based on Stockwell transform and Transformer. First, the S-transform is applied to the original EEG segments, acquiring accurate time-frequency representations. Subsequently, the obtained time-frequency matrices are grouped into different EEG rhythm blocks and compressed as vectors in these EEG sub-bands. After that, these feature vectors are fed into the Transformer network for feature selection and classification. Moreover, a series of post-processing methods were introduced to enhance the efficiency of the system. When evaluating the public CHB-MIT database, the proposed algorithm achieved an accuracy of 96.15%, a sensitivity of 96.11%, a specificity of 96.38%, a precision of 96.33%, and an area under the curve (AUC) of 0.98 in segment-based experiments, along with a sensitivity of 96.57%, a false detection rate of 0.38/h, and a delay of 20.62 s in event-based experiments. These outstanding results demonstrate the feasibility of implementing this seizure detection method in future clinical applications.

Keywords: automatic seizure detection; transformer; stockwell transform; EEG

Citation: Zhong, X.; Liu, G.; Dong, X.; Li, C.; Li, H.; Cui, H.; Zhou, W. Automatic Seizure Detection Based on Stockwell Transform and Transformer. *Sensors* **2024**, *24*, 77. https://doi.org/10.3390/s24010077

Academic Editor: Ting Li

Received: 21 October 2023
Revised: 15 December 2023
Accepted: 20 December 2023
Published: 22 December 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Epilepsy, caused by abnormal discharge of brain neurons, affects more than 50 million people worldwide [1]. Epilepsy is characterized by recurrent and sudden seizures, which may cause temporary loss of consciousness or perception, and involuntary body convulsions. Persistent and recurrent seizures can greatly disturb the patients' life and even endanger their safety. As a fundamental tool for studying the human brain, electroencephalogram (EEG) has become an important tool for assisting the clinical diagnosis of neurological diseases [2–5]. Presently, epileptic seizure events are mainly annotated by neurology experts based on clinical experience through analyzing long-term EEG recordings, which is time-consuming and laborious. Therefore, the development of automatic seizure detection systems, which can reduce the burden on medical staff and assist in patient treatment, has become a valuable research topic.

The research on automatic seizure detection has a history of several decades, and many promising results and preliminary applications have been achieved. One of the earliest seizure detection systems was proposed by Gotman [6] in the early 1980s. He extracted slope, rhythmicity, and sharpness as classification features from the brainwave signals decomposed into half-waves. Later, Gotman [7] and Qu [8] improved the method by developing a patient-specific false alarm model. Subsequently, many time-domain [9–12], frequency-do-

main [13–18], and deep learning methods [19–21] have been developed for seizure detection. For example, Acharya et al. [22] applied convolutional neural network (CNN) for the identification of epileptic EEG signals. Dong et al. [23] proposed an attention-based graph residual network with a redesigned focal loss function to address the class imbalance issue in epileptic seizure detection tasks. In the study of Tsiouris et al. [24], the LSTM model was used to classify EEG features extracted in time and frequency domains.

As EEG signals are typical non-stationary time-series signals, time-frequency analysis approaches such as Short-Time Fourier Transform, Wavelet Transform, and Empirical Mode Decomposition have been commonly employed to generate time-frequency representations for EEG signals [25–27]. Stockwell transform (S-transform), proposed by Stockwell et al. [28], is a combined approach of short-time Fourier transform and wavelet transform, allowing for multi-resolution analysis of time series with relatively low computational complexity. S-transform has been widely applied in various fields such as cardiac sound segmentation [29], power quality analysis [30–32], medical imaging [33], etc. Recently, researchers attempted to combine the S-transform with traditional classifiers and deep learning-based models for seizure detection, showing its effectiveness in analyzing epileptic EEG signals [34–36]. Therefore, in this study, the S-transform was adopted for accurate time-frequency representation of EEG signals.

The Transformer model with self-attention mechanisms was initially designed for machine translation [37]. Currently, it is widely used not only in natural language processing but also in areas such as computer vision [38], speech recognition [39], and motion imaging [40]. Multi-channel EEG signals are typical time series signals and also can be seen as an image, making them suitable for processing with Transformer models. Sun et al. [41] conducted experiments combining Transformer and 3D convolutional neural networks on three emotional EEG datasets, achieving better emotion recognition accuracy than other methods. Yan et al. [42] presented a model combining short-time Fourier transform and Transformer, demonstrating that their model can effectively utilize the time, frequency, and channel information in EEG signals to improve seizure prediction accuracy. Li et al. [43] introduced a novel graph neural network called the spatial-temporal graph attention network with a Transformer encoder (STGATE) for learning graph representations of emotion EEG signals and improving emotion recognition performance. The above studies indicate that Transformer has potential capabilities in EEG signal classification tasks.

This work proposes an effective method for seizure detection by the combination of S-transform and Transformer. Compared with short-time Fourier transform (STFT) and wavelet transform (WT), S-transform has the advantage of STFT and WT while maintaining lower computational complexity. In this work, the time-frequency matrices obtained by S-transform are compressed in specific frequency bands and then inputted into Transformer for automatic feature selection and classification. Transformer-based methods improve performance by assigning different weights to each channel of EEG signals while also increasing the interpretability of the model. The proposed Transformer contributes to improve performance by assigning different weights to each EEG channel while also increasing the interpretability of the model. The performance of the proposed approach is evaluated on the CHB-MIT epileptic EEG database. To the best of our knowledge, this is the first attempt in which S-transform and Transformer have been combined for seizure detection. Experimental results demonstrate the effectiveness of the proposed algorithm.

The rest of the article is organized as follows. Section 2 introduces the method for epileptic seizure detection, which includes S-transform, Transformer, and post-processing. Section 3 describes the CHB-MIT scalp epileptic EEG dataset and experimental results based on segment-level and event-level. Section 4 is devoted to discussing the results and comparing the performance with other algorithms. Finally, Section 5 presents the conclusion.

2. Methods

Figure 1 shows the overall workflow of the proposed seizure detection method, which mainly consists of three essential parts: pre-processing (segmentation and S-transform), Transformer, and multi-layer perception. In this work, the multi-channel EEG recordings were divided into 4-s (1024-point) segments.

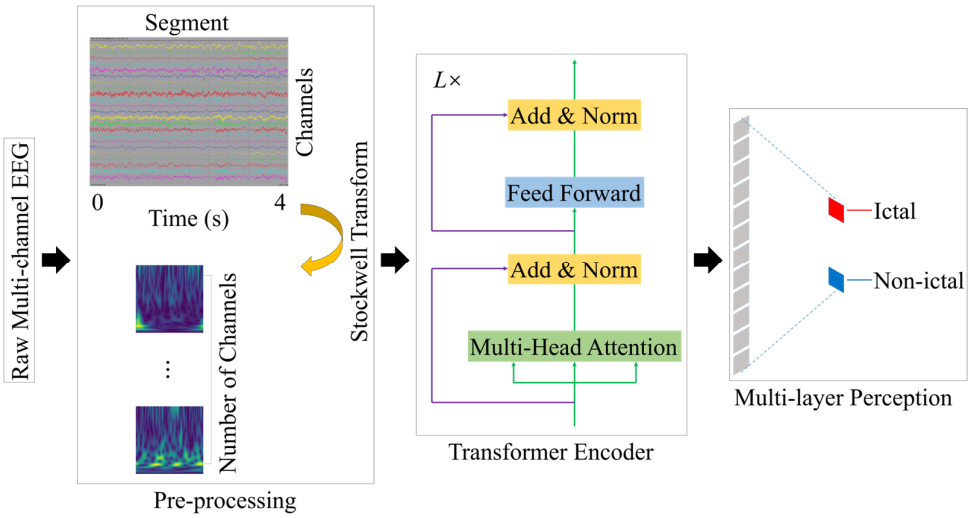


Figure 1. The workflow of the proposed method for seizure detection.

2.1. Stockwell Transform

S-transform is a time-frequency domain analysis method proposed by geophysicist Stockwell [28] in 1996. By combining the advantages of short-time Fourier transform (STFT) and wavelet transform (WT), it has become an effective tool for analyzing and processing non-stationary EEG signals. The S-transform spectrogram $S_x(\tau, f)$ of time domain signal $x(t)$ is defined by:

$$S_x(\tau, f) = e^{-i2\pi f\tau} W_x(\tau, d) \tag{1}$$

$$W_x(\tau, d) = \int_{-\infty}^{+\infty} x(t)\omega(t-\tau, d)dt \tag{2}$$

where $W_x(\tau, d)$ denotes the wavelet transform of $x(t)$ and $\omega(t, f)$ is the mother wavelet, which is defined as:

$$\omega(t, f) = \frac{|f|}{\sqrt{2\pi}} e^{\frac{-t^2 f^2}{2}} e^{-i2\pi ft} \tag{3}$$

Ultimately, the S-transform can be given as follows:

$$S_x(\tau, f) = \int_{-\infty}^{+\infty} x(t) \frac{|f|}{\sqrt{2\pi}} e^{\frac{-(\tau-t)^2 f^2}{2}} e^{-i2\pi ft} dt \tag{4}$$

where the $x(t)$ represents the segmented 4-s EEG signals in this study. Each EEG segment processed by S-transform returns a time-frequency matrix with size of 128 × 1024, where 128 represents the frequency range from 1 to 128 Hz, and 1024 expresses the time points. Figure 2 shows a 4-s segment of typical non-ictal EEG and a 4 s segment of typical epileptic EEG selected from patient 5, along with the corresponding S-transform spectrograms. It is evident from the figure that not only is the amplitude of the epileptic EEG significantly higher than the non-ictal EEG, but there is also a notable difference in energy between the two signals in the frequency range of 20 to 50 Hz.

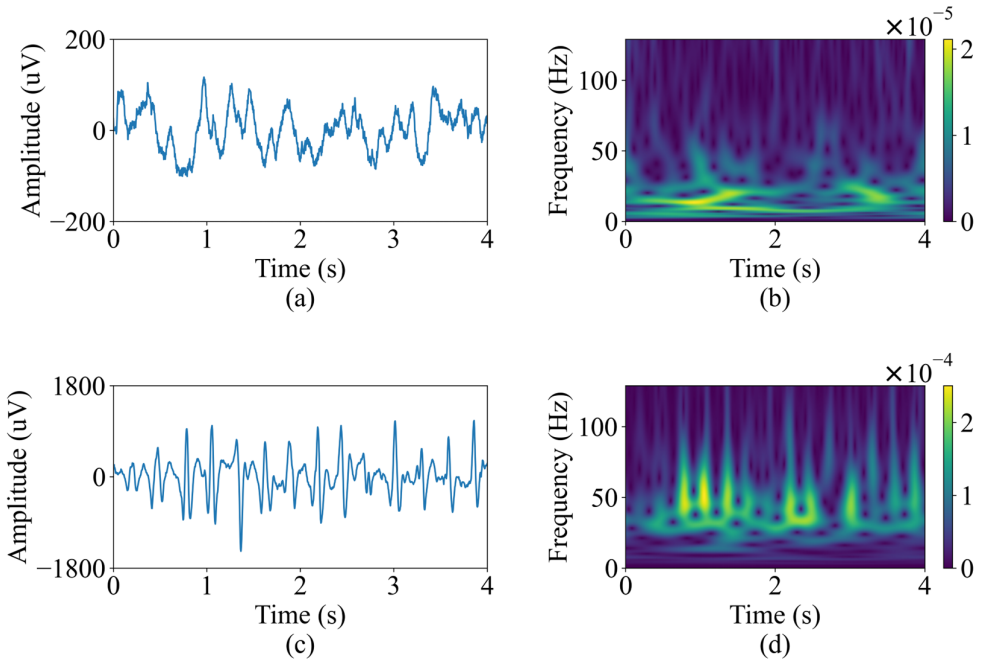


Figure 2. Non-seizure and seizure EEG signals and their S-transform spectrograms. (**a**) Non-ictal EEG signal. (**b**) S-transform of non-ictal EEG. (**c**) Ictal EEG signal. (**d**) S-transform of ictal EEG.

Considering that epileptic EEG signals are concentrated in the frequency range of 3 to 30 Hz [35], the frequency range of 1–50 Hz is selected in order to eliminate power frequency interference, and then divided into 6 sub-bands: delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), gamma1 (30–40 Hz), and gamma2 (40–50 Hz). For a 4-s EEG segment, the 4-s time axis is partitioned into two parts with a 2-s interval. Hence, the time-frequency matrix obtained with S-transform within 1–50 Hz is divided into 12 sub units. The summation of the squared moduli of S-transform in each unit are sequentially concatenated to obtain a feature map of size $n \times 12$ as input to the model, where n represents the number of channels. Figure 3 depicts the above time-frequency compression process for a single channel.

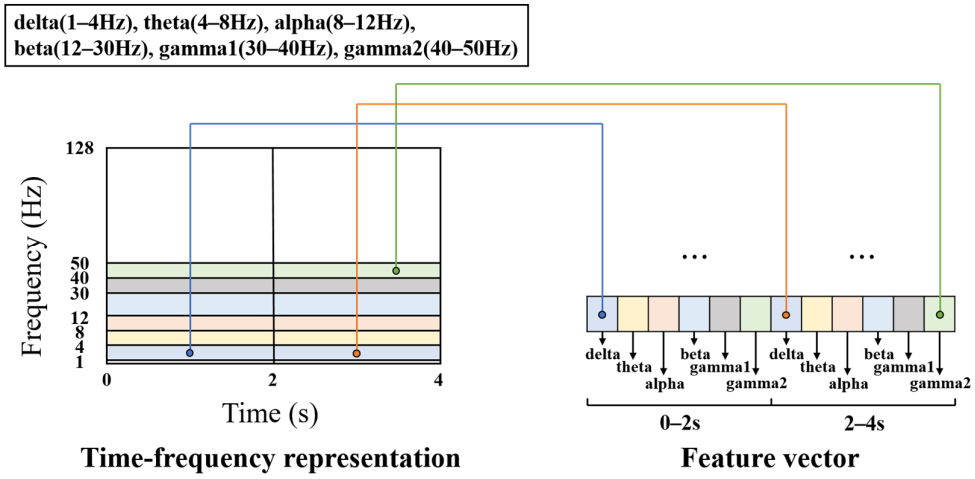


Figure 3. Schematic diagram for compression of EEG time-frequency representation obtained with S-transform in a single channel. Different colors represent corresponding EEG rhythms, including delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), gamma1 (30–40 Hz) and gamma2 (40–50 Hz).

2.2. Transformer

The Transformer network was originally designed for machine translation, and the sequences fed into the network need to be pre-embedded to obtain a matrix of shape number of words × embedding dimension. We regard each channel of the EEG recording as a word in a sequence. Therefore, the above process is also known as "channel embedding" [44]. In this study, since we want to keep the channel order of each EEG recording consistent, no position encoding was performed on channels.

Given that the goal of epilepsy detection task was actually a classification task, we only used the encoder module of the Transformer, which was stacked L times. The S-transformed and compressed EEG signal S serves as the input feature map of the Transformer encoder. The Transformer encoder consists of two parts: Multi-head Self-Attention (MSA) and Multi-Layer Perception (MLP). Both parts use layer normalization, and their outputs adopt residual connection structures. The self-attention mechanism can be described as a mapping from a query matrix (Q) to a set of key (K)-value (V) pairs. Q and K have a dimension of d_k, and V has a dimension of d_v. The output of the sequence S after the self-attention mechanism can be calculated by the following equation:

$$[Q, \quad K, \quad V] = S[W^Q, \quad W^K, \quad W^V] \tag{5}$$

$$\mathrm{SA}(S) = softmax(QK^T/\sqrt{d_k})V \tag{6}$$

where $W^Q \in R^{d_{model} \times d_k}$, $W^K \in R^{d_{model} \times d_k}$, and $W^V \in R^{d_{model} \times d_v}$ are the linearity transformation matrices. In MSA mechanism, multiple self-attention operations are run in parallel, and their concatenate outputs are returned.

$$\mathrm{MSA}(S) = concat(\mathrm{SA}_1(S), \cdots, \mathrm{SA}_h(S))W^O \tag{7}$$

where the coefficient matrix $W^O \in R^{h \times d_v \times d_{model}}$. In this research, $d_k = d_v = d_{model}/h = 4$, and we utilize $h = 3$ concurrent attention heads. The computation process of the aforementioned multi-head self-attention mechanism is depicted in Figure 4.

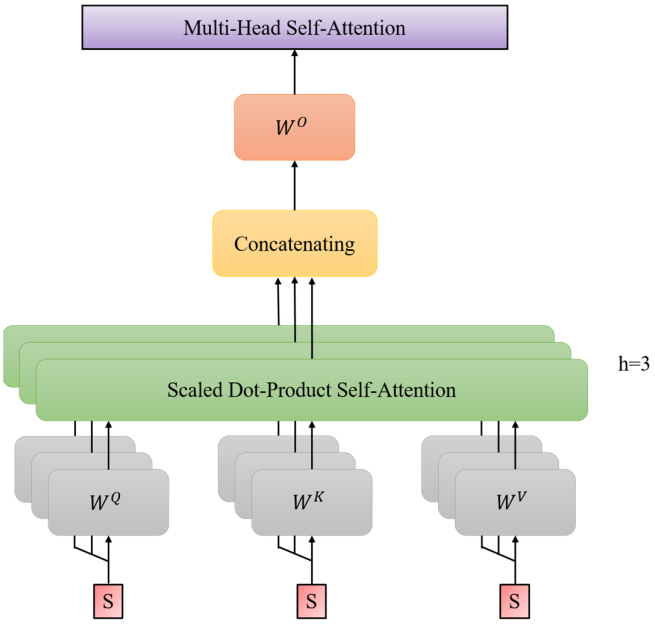


Figure 4. Structure diagram of multi-head self-attention mechanism.

The Transformer encoder used in this work consists of L layers, and the output of each layer serves as the input to the next layer. This process can be expressed by the following two equations:

$$y_i = \mathrm{MSA}(\mathrm{LN}(S_{i\text{-}1})) \tag{8}$$

$$S_i = S_{i\text{-}1} + y_i + \mathrm{MLP}(\mathrm{LN}(y_i)) \tag{9}$$

where $i = 1, \cdots, L$ and $L = 6$ is chosen.

Finally, the max-pooling operation is performed along the first dimension of S to obtain the input of the MLP module, and pass the result through a softmax layer to output the probabilities for seizure and non-seizure.

2.3. Post-Processing

When using long-term continuous EEG recordings for event-based assessment, isolated false detections are often encountered. We apply moving average filtering (MAF), collar technique, and K-of-N method to reduce the false detection rate (FDR) of the algorithm. MAF is applied before thresholding to smooth the predicted scores. The collar technique is mainly used to prevent seizure segments that are correctly detected from being filtered out. In this study, K and N are set to 5 and 10, respectively, and a 40 s window consisting of N epochs is used to slide over the model prediction results. If five or more samples in the window are judged as seizures, the time span is considered a seizure period. The post-processing procedure is illustrated in Figure 5.

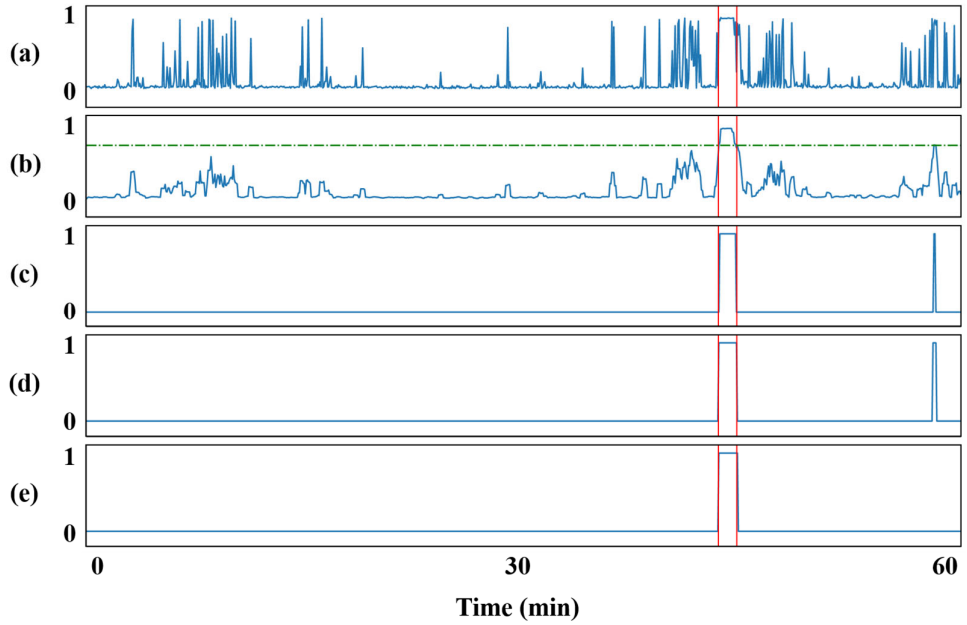


Figure 5. The post-processing procedure of 1-h EEG data. (**a**) The model prediction. (**b**) The outputs processed by MAF. (**c**) The binary values obtained from threshold determination. (**d**) The results after collar technology. (**e**) The final decisions after K-of-N discrimination. The vertical red lines represent expert-annotated seizure events, while the horizontal green line represents the threshold set during the binarization operation.

3. Experiments and Results

3.1. EEG Dataset

The long-term scalp EEG dataset used in this study was collected at the Children's Hospital Boston, and consists of EEG recordings from pediatric subjects with intractable seizures [45,46]. The recordings, grouped into 24 cases, were collected from 23 subjects (5 males, ages 3–22, and 17 females, ages 1.5–19) [47]. All signals were sampled at 256 Hz with 16-bit resolution. Most files contain 23 EEG channels (24 or 26 in a few cases). The international 10–

20 system of EEG electrode positions and nomenclature was used for these recordings. In summary, these records include 182 seizures (166 in the original set of 24 cases).

The details of the used CHB-MIT dataset are listed in Table 1. In this work, we evaluated the proposed model based on the segments for all patients. However, due to the fact that the duration of each seizure in chb16 is less than 15 s, this patient was excluded from the event-based evaluation of the model.

Table 1. Details of the used CHB-MIT EEG dataset.

Patient	Gender	Age (Years)	No. of Used Channels	No. of Epileptic Events	Duration of Epileptic Seizures (s)
1	F	11	23	7	442
2	M	11	23	3	172
3	F	14	23	7	402
4	M	22	23	4	378
5	F	7	23	5	558
6	F	1.5	23	10	138
7	F	14.5	23	3	325
8	M	3.5	23	5	919
9	F	10	23	4	276
10	M	3	23	7	447
11	F	12	23	3	806
12	F	2	23	40	1475
13	F	3	18	12	535
14	F	9	23	8	109
15	M	16	18	20	1992
16	F	7	18	10	84
17	F	12	23	3	293
18	F	18	23	6	317
19	F	19	23	3	236
20	F	6	23	8	294
21	F	13	23	4	199
22	F	9	23	3	204
23	F	6	23	7	424
24	-	-	23	16	511

3.2. Experimental Process and Evaluation

In the segment-based experiments, each patient's ictal EEG recordings were divided in 4 s seizure segments with a sliding window based on the start and end times of the seizures annotated by the experts, and an equal number of 4 s normal EEG segments were randomly selected. Since seizure EEG data is much less than normal EEG data and in order to enhance the system's generalization ability, we used 50% overlapping sliding window when dividing the ictal data, while no overlap strategy was performed when dividing normal EEG data. After segmenting the dataset, we generated random seeds to shuffle the segmented dataset during model training. The dataset was then grouped into training and testing subsets with a 3:1 ratio.

For the segment-based evaluation, five metrics were introduced to assess the performance of the model: accuracy, sensitivity, specificity, precision, and area under the curve (AUC).

$$Accuracy = \frac{TN + TP}{TN + FP + FN + TP} \tag{10}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{11}$$

$$Specificity = \frac{TN}{TN + FP} \tag{12}$$

$$Precision = \frac{TP}{FP + TP} \tag{13}$$

where, *TP* (true positive) and *TN* (true negative) refer to the number of ictal and non-ictal segments that are correctly recognized by our detection method, respectively. *FP* (false positive) denotes the number of non-ictal EEG segments incorrectly judged as ictal by the detection method, and *FN* (false negative) means that the ictal segment is labeled to be the non-ictal segment.

AUC stands for the area under the curve of the receiver operating characteristic curve (ROC), which denotes the probability that the positive samples are assigned a higher score than the negative samples [48].

In the event-based experiments, partial seizure events of each patient were used as training samples. The processing of the training set followed a similar approach to the segment-based processing, with the exception that the optimal model parameters were saved at the end. Thereafter, all recordings of the patient except for those training seizure events are used for testing. All the test files are arranged in chronological order according to the recorded time, and the onset as well as offset times of epileptic seizures are annotated based on the instruction files. After post-processing operations, the previously saved model's output is considered a correct detection if it predicts a seizure within the onset and offset time range. Conversely, if the model predicts a seizure outside the range, it is considered a false alarm. CHB-MIT dataset A total of 865.15 h of EEG data containing 97 seizures from the CHB-MIT dataset were utilized for event-based performance testing, alone with other 64 seizures used for model training.

For the event-based evaluation, three measures are utilized to evaluate the model in clinical practice: sensitivity, FDR, and detection delay. Sensitivity is computed by dividing the number of correctly detections by the number of testing seizures for per patient. FDR represents the number of falsely detected seizures by the model within one hour. Detection delay is the time interval between the point at which the model makes a correct seizure detection and the expert-annotated onset time of the seizure.

3.3. Results

The model was implemented using PyTorch 1.13.1 in Python 3.9. All the results presented below were obtained through experiments conducted on a GPU configured with GeForce RTX 3050 from United States-based company NVIDIA.

Table 2 lists the segment-based experimental results. On average, the accuracy of 96.15%, the sensitivity of 96.11%, the specificity of 96.38%, the precision of 96.33%, and the AUC of 0.98 are achieved. There are more than half of the patients having an AUC above 0.99, while 11 patients have a sensitivity greater than 97%. Patients 15 and 21 have relatively low classification accuracies, both falling below 90%.

Table 2. Detection results of the proposed method on segment-based metrics.

Patient	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	AUC
1	99.11	98.21	100	100	1
2	97.73	95.24	100	100	0.9917
3	95.05	96	94.12	94.12	0.9694
4	96.88	97.56	96.36	95.24	0.9849
5	99.29	98.73	100	100	1

6	97.22	94.44	100	100	0.9938
7	96.3	95.24	97.44	97.56	0.9939
8	97.40	97.30	97.50	97.30	0.9877
9	100	100	100	100	1
10	98.25	98.39	98.08	98.39	0.9932
11	96.53	97.73	95.61	94.51	0.9929
12	97.10	96.00	98.28	98.36	0.9919
13	96.92	95.45	98.44	98.44	0.9744
14	97.78	100	96.30	94.74	0.9979
15	89.02	84.21	93.88	93.27	0.9377
16	100	100	100	100	1
17	93.06	94.59	91.43	92.11	0.9521
18	93.67	90.48	97.30	97.44	0.9665
19	96.61	100	93.75	93.10	0.9954
20	94.44	92.68	96.77	97.44	0.9929
21	83.67	95.65	73.08	75.86	0.8094
22	96.08	95.65	96.43	95.65	0.9984
23	97.14	94.55	100	100	0.9975
24	98.39	98.44	98.33	98.44	0.9982
Total	96.15	96.11	96.38	96.33	0.98

The event-based experimental results are shown in Table 3. For event-based evaluation, we obtained an average sensitivity of 96.57%, an average FDR of 0.38/h, and an average detection delay of 20.62 s. In addition, apart from patients 12, 15, 23, and 24, no seizure events were missed in the other patients. As for patient 16, the average duration of her epileptic seizures was only 8.4 s. Even with 50% overlapping sampling of 4 s segments, the training data were severely insufficient. Therefore, this patient was excluded from the event-based experiments in this study.

Table 3. Detection results of the proposed method on event-based metrics.

Patient	Test Set Duration (h)	No. of Training Seizures	No. of Testing Seizures	No. of True Detections	Sensitivity (%)	FDR (%)	Latency (s)
1	37.38	3	4	4	100	0.11	19
2	34	2	1	1	100	0	36
3	34	4	3	3	100	0.56	16.3
4	149.38	2	2	2	100	0.39	9.5
5	37	2	3	3	100	0	33.3
6	48.52	4	3	3	100	1.26	6
7	63.05	1	2	2	100	0.03	19.5
8	19	1	4	4	100	0.58	64.25
9	62.27	2	2	2	100	0.61	20.5
10	44	3	4	4	100	0	8.75
11	33.79	1	2	2	100	0.06	4
12	16.67	4	15	12	80	0.42	3.08
13	29	4	7	7	100	0.86	19.14
14	22	4	4	4	100	0.05	9.5
15	34	6	14	11	78.57	0.71	21.09
17	19	2	2	2	100	0.05	41.5
18	31.63	4	3	3	100	0.41	15
19	27.93	2	2	2	100	0.11	11
20	24.63	3	4	4	100	0.77	6.75

21	31	2	2	2	100	0.97	37.25
22	30	1	2	2	100	0.17	23
23	21.6	1	4	3	75	0.69	31.33
24	15.3	6	8	7	87.5	0	18.43
Total	865.15	64	97	89	96.57	0.38	20.62

4. Discussion

4.1. Comparison with Existing Methods

Table 4 lists some state of the art seizure detection methods that have also been evaluated on the CHB-MIT EEG database. Ansari et al. [49] achieved a sensitivity of 85%, specificity of 89.06%, and classification accuracy of 89.06% on the CHB-MIT dataset by combining frequency domain features of EEG signals generated by wavelet packet decomposition with a neutrosophic logic-based k-means nearest neighbor (NL-k-NN) classifier. Janjarasjitt [50] extracted wavelet features from scalp EEG recordings and classified them using support vector machines (SVM), achieving an accuracy of 96.87%, sensitivity of 72.99%, and specificity of 98.13%. He et al. [51] used graph attention networks (GAT) and BiLSTM as the front-end for extracting spatial features and the back-end for exploring temporal relationships. Through extensive experiments, they demonstrated that this model can effectively detect epileptic seizures from raw EEG signals. The automatic seizure detection system proposed by Yao et al. [52], based on transfer learning of VGGNet-16 and gated recurrent unit (GRU), achieved sensitivity, specificity, and accuracy of 90.12%, 96.32%, and 96.31%, respectively. However, their experiments were conducted only on 12 patients. Cura et al. [53] computed features like higher-order joint time-frequency (HOJ-TF) moments and gray-level co-occurrence matrix (GLCM) through synchrosqueezing transform (SST) to obtain high-resolution time-frequency representations of EEG signals. These representations were combined with machine learning algorithms achieved a promising classification performance. Hu et al. [54] introduced local mean decomposition (LMD) and feature extraction processes to reduce computational complexity while maintaining the non-stationarity of EEG signals. They used BiLSTM to achieve a sensitivity of 93.61% and specificity of 91.85%. Duan et al. [55] proposed an epileptic seizure detection method based on deep metric learning, with an average accuracy of 86.68% and average specificity of 93.71% on the CHB-MIT dataset. Shyu et al. [56] presented an end-to-end deep learning model comprising an inception module and a residual module for seizure detection. Although their method achieved higher accuracy of 98.34% and specificity of 98.79% on the CHB-MIT database, it exhibited much lower sensitivity of 73.08%. Jiang et al. [57] used a seizure detection method based on the brain functional network structure and time-frequency multi-domain features. They employed SVM classifier for ictal EEG classification. However, their method operated on multi-domain hand-crafted features and irrelevant features needs to be eliminated using principal component analysis (PCA), which increase the complexity of the algorithm. Considering that seizure episodes have much shorter durations compared to non-seizure EEG, Gao et al. [58] utilized generative adversarial network (GAN) for data augmentation, and used one-dimensional convolutional neural network (1DCNN) for seizure detection, achieving a sensitivity of 93.53% and specificity of 99.05%. Their achieved overall sensitivity is lower than our method.

Most of the aforementioned studies only employed segmented EEG for evaluation. Event-based assessments for epileptic seizure detection are more concordant with practical clinical applications and proved to be challenging due to the frequent appearance of artifacts in long-term continuous EEG. Zhang et al. [59] combined wavelet transform and bidirectional gated recurrent unit (Bi-GRU) network followed by certain post-processing steps, achieving an average sensitivity of 93.89% and an average specificity of 98.49%. Among the 128 seizure events used, the model only missed four detections and reduced the false alarm rate to 0.31 per hour, indicating the potential superiority of the Bi-GRU network in long-term EEG applications. Yoshiba et al. [60] yielded a detection delay of 7.39 s by using a single EEG channel combined with pretrained ResNet. However, this study only used data

from 10 patients (3–19 years old). Samiee et al. [61] proposed a feature extraction method based on sparse rational decomposition and Local Gabor Binary Patterns (LGBP), with a sensitivity of 91.13% and FDR of 0.35/h at event-based level and a delay of 5.98 s. Compared with the above research, our proposed method obtained the highest event-based sensitivity within a competitive FDR.

In previous studies, CNNs have been employed for encoding and classifying EEG features. For instance, Sun et al. [62] proposed a subject transfer neural network (STNN) by integrating CNN with self-attention, achieving satisfactory results in motor imagery classification tasks. However, the local convolutional structure in CNNs poses difficulty to capture the global features of input signals. Meanwhile, CNNs require serial operations at each time step, resulting in lower computational efficiency when handling long-term time-series data. In contrast, the Transformer encoder with multi-head self-attention utilized in this work can capture long-range correlations and global features of input signals, and allows for parallel computation and better classification capability. Moreover, in comparison with the continuous wavelet transform (CWT) used by Sun et al. [62], the S-transform adopted in this method combines the advantages of CWT and Short-Time Fourier Transform (STFT) while maintaining lower computational complexity, enabling better extraction of time-frequency features from EEG signals.

Overall, the performance and stability of the proposed method are satisfactory. These results verify the effectiveness of the combination of S-transform and Transformer in epileptic seizure detection.

Table 4. Performance comparison of different seizure detection methods reported for CHB-MIT dataset.

Author	Method	Segment-Based					Event-Based		
		Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	AUC (%)	Sensitivity (%)	FDR (/h)	Latency (s)
Ansari et al. [49]	kNN	89.06	85	89.06	-	-	-	-	-
Janjarasjitt et al. [50]	Wavelet + SVM	96.87	72.99	98.13	-	-	-	-	-
He et al. [51]	GAT + BiLSTM	98.52	97.75	94.34	-	96.81	-	-	-
Yao et al. [52]	Transfer learning + GRU	96.31	90.12	96.32	-	-	-	-	-
Cura et al. [53]	SST + kNN	95.1	90.3	-	93.4	-	-	-	-
Hu et al. [54]	LMD + BiLSTM	-	93.61	91.85	-	-	-	-	-
Duan et al. [55]	Deep metric learning	86.68	79.64	93.71	-	-	-	-	-
Shyu et al. [56]	Inception and Residual model	98.34	73.08	98.79	-	-	-	-	-
Jiang et al. [57]	PMNet + SVM	96.67	97.72	95.62	-	-	-	-	-
Gao et al. [58]	GAN + 1DCNN	93.53	99.05	-	-	-	-	-	-
Zhang et al. [59]	Bi-GRU	98.49	93.89	98.49	-	-	95.49	0.31	-
Yoshiba et al. [60]	ResNet	-	88.73	98.98	-	-	-	-	7.39
Samiee et al. [61]	Sparse rational decomposition + LGBP	-	70.40	99.10	-	-	91.13	0.35	5.98
This work	S-transform + Transformer	96.15	96.11	96.38	96.33	98	96.57	0.38	20.62

4.2. Visualization of t-Distributed Stochastic Neighbor Embedding (t-SNE)

Figure 6 illustrates the visualization of the sample distribution obtained from seizure and non-seizure samples extracted from the EEG recordings of three patients (chb08, chb11, chb15) using the *t*-SNE algorithm. The upper three plots (a), (b), and (c) depict the two-dimensional projection of the 4 s original EEG segments from the aforementioned three patients, while the bottom three plots (d), (e), and (f) show the corresponding sample distributions in the time-frequency domain after applying the S-transform. In these plots, the red points represent normal samples while the blue points represent epileptic samples. It is evident from Figure 6c,f that the EEG time-frequency features obtained by the S-transform exhibit better separability when compared to the scattered distribution of the original EEG samples embedded in time domain. These observations indicate the effectiveness of the S-transform in assisting to feature extraction from normal and epileptic EEGs. However, these time-frequency features are not completely separable, necessitating further extraction and classification with the assistance of the Transformer Encoder. To achieve better seizure detection performance, we incorporated S-transform and Transformer in this work.

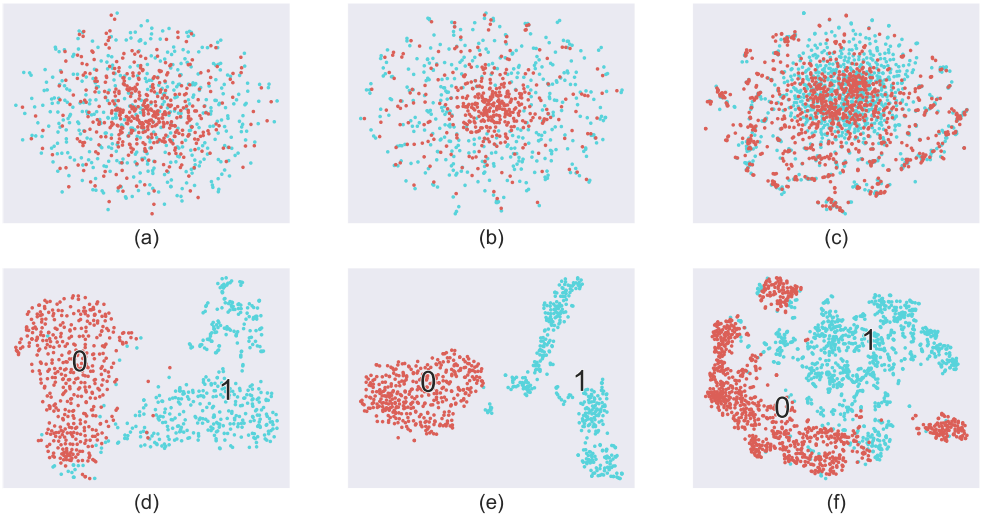


Figure 6. *t*-SNE visualization of samples from chb08, chb11, and chb15. (**a**–**c**) Distributions of original EEG samples from chb08, chb11, and chb15. (**d**–**f**) Corresponding distributions of those EEG samples after S-transform. The red points labeled with the number 0 represent normal samples while the blue points labeled with the number 1 represent epileptic samples.

4.3. Attention to EEG Channels

In recent years, some studies have aimed not only to achieve high classification accuracy and sensitivity but also to determine which channels of multi-channel EEG recordings related to seizure onsets [63]. In the proposed Transformer encoder, the self-attention module can characterize the active channels by quantitating EEG channel attention weights. Figure 7 illustrates the attention weight matrices generated by the last encoder layer of the proposed model for each of the three attention heads on patient chb23, along with their sum. The final output of the Transformer encoder is the product of the weight matrix (W) and the input (S). By summing and averaging the elements in each column vector in W, the channel attention weight vector can be obtained. As shown in Figure 8, the model assigns higher attention to channels FP1-F3, F3-C3, C3-P3, FZ-CZ, CZ-PZ, T7-FT9, and FT10-T8, indicating that the epileptic EEG activity is likely to be active in these channels.

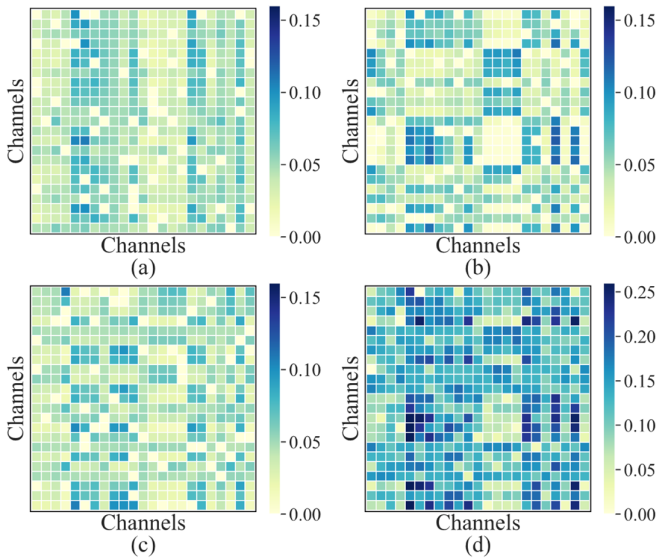


Figure 7. The visualization of attention mechanism. (**a**–**c**) The channel weight matrix generated by the multi-head attention of the last encoder layer of Transformer. (**d**) The sum of the above three.

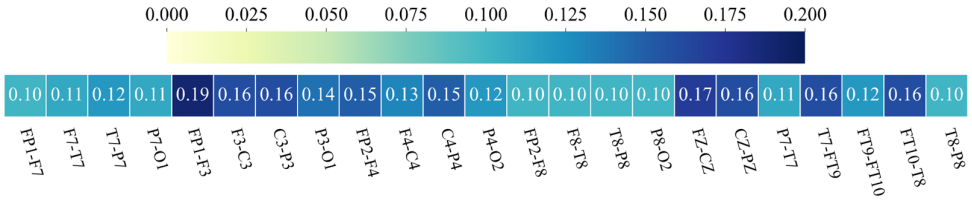


Figure 8. The attention weights assigned to each EEG channel for patient 23.

4.4. Future Work

Due to the significant individual variability observed in the severity and duration of epileptic seizures, extensive research of epileptic detection, including the present study, has focused on individualized patient-specific methods. However, a critical avenue for future research lies in enhancing the generalization and capabilities of the model to make it more suitable for real-world clinical application. To this end, exploration into cross-subject epileptic detection is in critical need. Our future endeavors will involve conducting more experiments to further verify the generalization performance of the proposed model for patient-independent seizure detection, extending its applicability even in across-dataset scenarios.

5. Conclusions

In this study, we propose a novel automatic seizure detection approach based on S-transform and Transformer. The S-transform enables more comprehensive time-frequency representations compared to STFT and wavelet transform, facilitating the Transformer encoder to learn more distinctive features. Meanwhile, the proposed Transformer model can assign unequal attention weights to different EEG channels, thereby extracting spatial features of multi-channel EEG signals and enhancing the interpretability of the model by preserving the original labels of the channels. The method has been investigated on the CHB-MIT database and achieves 96.15% accuracy, 96.11% sensitivity, 96.38% specificity, 96.33% precision, and 0.98 AUC on segment-based evaluation. Additionally, sensitivity of 96.57%, 0.38/h FDR, and 20.62 s average latency are yielded under event-based level. These outstanding results indicate the feasibility of implementing this seizure detection method in clinical applications.

Author Contributions: Conceptualization, X.Z. and W.Z.; methodology, X.Z.; software, X.Z. and X.D.; validation, X.Z., X.D. and C.L.; formal analysis, C.L., H.L. and H.C.; investigation, X.Z., X.D. and H.L.; resources, X.D. and G.L.; data curation, X.Z., C.L. and H.C.; writing—original draft preparation, X.Z.; writing—review and editing, G.L. and W.Z.; visualization, X.Z. and W.Z.; supervision, W.Z.; project administration, W.Z.; funding acquisition, W.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (No. 62271291), the Key Program of the Natural Science Foundation of Shandong Province (No. ZR2020LZH009), and the Shenzhen Science and Technology Program (GJHZ20220913142607013).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The CHB-MIT Database analyzed in this study is available from https://physionet.org/content/chbmit/1.0.0/.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Epilepsy. Available online: https://www.who.int/zh/news-room/fact-sheets/detail/epilepsy (accessed on 13 October 2023).
2. Xin, Q.; Hu, S.; Liu, S.; Zhao, L.; Zhang, Y.D. An Attention-Based Wavelet Convolution Neural Network for Epilepsy EEG Classification. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2022**, *30*, 957–966.
3. Patil, A.U.; Dube, A.; Jain, R.K.; Jindal, G.D.; Madathil, D. Classification and Comparative Analysis of Control and Migraine Subjects Using EEG Signals. In *Information Systems Design and Intelligent Applications: Proceedings of Fifth International Conference INDIA 2018. Advances in Intelligent Systems and Computing (AISC 862)*; Springer: Singapore, 2019; pp. 31–39.
4. Smith, S.J. EEG in the diagnosis, classification, and management of patients with epilepsy. *J. Neurol. Neurosurg. Psychiatry* **2005,** *76* (Suppl. 2), ii2–ii7.
5. Xin, Q.; Hu, S.H.; Liu, S.Q.; Ma, X.L.; Lv, H.; Zhang, Y.D. Epilepsy EEG Classification Based on Convolution Support Vector Machine. *J. Med. Imaging Health Inform.* **2021**, *11*, 25–32.
6. Gotman, J. Automatic Recognition of Epileptic Seizures in the EEG. *Electroen. Clin. Neurophysiol.* **1982**, *54*, 530–540.
7. Gotman, J. Automatic Seizure Detection—Improvements and Evaluation. *Electroen. Clin. Neurophysiol.* **1990**, *76*, 317–324.
8. Qu, H.; Gotman, J. Improvement in Seizure Detection Performance by Automatic Adaptation to the Eeg of Each Patient. *Electroen. Clin. Neuro.* **1993**, *86*, 79–87.
9. Kabir, E.; Siuly; Cao, J.L.; Wang, H. A computer aided analysis scheme for detecting epileptic seizure from EEG data. *Int. J. Comput. Int. Syst* **2018**, *11*, 663–671.
10. Samiee, K.; Kovacs, P.; Gabbouj, M. Epileptic Seizure Classification of EEG Time-Series Using Rational Discrete Short-Time Fourier Transform. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 541–552.
11. Tzallas, A.T.; Tsipouras, M.G.; Fotiadis, D.I. Epileptic Seizure Detection in EEGs Using Time-Frequency Analysis. *IEEE Trans. Inf. Technol. Biomed.* **2009**, *13*, 703–710.
12. Logesparan, L.; Casson, A.J.; Rodriguez-Villegas, E. Optimal features for online seizure detection. *Med. Biol. Eng. Comput.* **2012**, *50*, 659–669.
13. Riaz, F.; Hassan, A.; Rehman, S.; Niazi, I.K.; Dremstrup, K. EMD-Based Temporal and Spectral Features for the Classification of EEG Signals Using Supervised Learning. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2016**, *24*, 28–35.
14. Al Ghayab, H.R.; Li, Y.; Siuly, S.; Abdulla, S. Epileptic EEG signal classification using optimum allocation based power spectral density estimation. *Iet. Signal Process* **2018**, *12*, 738–747.
15. Al Ghayab, H.R.; Li, Y.; Siuly, S.; Abdulla, S. Epileptic seizures detection in EEGs blending frequency domain with information gain technique. *Soft Comput.* **2019**, *23*, 227–239.
16. Khamis, H.; Mohamed, A.; Simpson, S. Frequency-moment signatures: A method for automated seizure detection from scalp EEG. *Clin. Neurophysiol.* **2013**, *124*, 2317–2327.
17. Rana, P.; Lipor, J.; Lee, H.; van Drongelen, W.; Kohrman, M.H.; Van Veen, B. Seizure Detection Using the Phase-Slope Index and Multichannel ECoG. *IEEE Trans. Biomed. Eng.* **2012**, *59*, 1125–1134.
18. Kapoor, B.; Nagpal, B.; Jain, P.K.; Abraham, A.; Gabralla, L.A. Epileptic Seizure Prediction Based on Hybrid Seek Optimization Tuned Ensemble Classifier Using EEG Signals. *Sensors* **2022**, *23*, 423.
19. Hussein, R.; Palangi, H.; Ward, R.K.; Wang, Z.J. Optimized deep neural network architecture for robust detection of epileptic seizures using EEG signals. *Clin. Neurophysiol.* **2019**, *130*, 25–37.
20. Liu, G.; Tian, L.; Zhou, W. Patient-Independent Seizure Detection Based on Channel-Perturbation Convolutional Neural Network and Bidirectional Long Short-Term Memory. *Int. J. Neural Syst.* **2022**, *32*, 2150051.
21. Li, Y.; Liu, Y.; Guo, Y.Z.; Liao, X.F.; Hu, B.; Yu, T. Spatio-Temporal-Spectral Hierarchical Graph Convolutional Network With Semisupervised Active Learning for Patient-Specific Seizure Prediction. *IEEE Trans. Cybern.* **2022**, *52*, 12189–12204.
22. Acharya, U.R.; Oh, S.L.; Hagiwara, Y.; Tan, J.H.; Adeli, H. Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Comput. Biol. Med.* **2018**, *100*, 270–278.
23. Dong, C.; Zhao, Y.; Zhang, G.; Xue, M.; Chu, D.; He, J.; Ge, X. Attention-based Graph ResNet with focal loss for epileptic seizure detection. *J. Ambient Intell. Smart Environ.* **2022**, *14*, 61–73.
24. Tsiouris, K. M.; Pezoulas, V.C.; Zervakis, M.; Konitsiotis, S.; Koutsouris, D.D.; Fotiadis, D.I. A Long Short-Term Memory deep learning network for the prediction of epileptic seizures using EEG signals. *Comput. Biol. Med.* **2018**, *99*, 24–37.
25. Truong, N.D.; Nguyen, A.D.; Kuhlmann, L.; Bonyadi, M.R.; Yang, J.; Ippolito, S.; Kavehei, O. Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. *Neural Netw.* **2018**, *105*, 104–111.
26. Alickovic, E.; Kevric, J.; Subasi, A. Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction. *Biomed. Signal Process. Control* **2018**, *39*, 94–102.
27. Molla, M.K.I.; Hassan, K.M.; Islam, M.R.; Tanaka, T. Graph Eigen Decomposition-Based Feature-Selection Method for Epileptic Seizure Detection Using Electroencephalography. *Sensors* **2020,** *20*, 4639.
28. Stockwell, R.G.; Mansinha, L.; Lowe, R.P. Localization of the complex spectrum: The S transform. *IEEE Trans. Signal Process.* **1996**, *44*, 998–1001.
29. Moukadem, A.; Dieterlen, A.; Hueber, N.; Brandt, C. A robust heart sounds segmentation module based on S-transform. *Biomed. Signal Process. Control* **2013**, *8*, 273–281.
30. Lee, C.Y.; Shen, Y.X. Feature Analysis of Power Quality Disturbance in Smart Grid Using S-Transform and TT-Transform. *Int. Rev. Electr. Eng.-I* **2012**, *7*, 4208–4220.

31. Raj, S.; Phani, T.C.K.; Dalei, J. Power Quality Analysis Using Modified S-Transform on ARM Processor. In Proceedings of the 2016 Sixth International Symposium on Embedded Computing and System Design (Ised 2016), Patna, India, 15–17 December 2016; pp. 166–170.
32. Dash, P.K.; Panigrahi, B.K.; Panda, G. Power quality analysis using S-Transform. *IEEE Trans. Power Deliv.* **2003**, *18*, 406–411.
33. Assous, S.; Humeau, A.; Tartas, M.; Abraham, P.; L'Huillier, J.P. S-transform applied to laser Doppler flowmetry reactive hyperemia signals. *IEEE Trans. Biomed. Eng.* **2006**, *53*, 1032–1037.
34. Liu, G.; Zhou, W.; Geng, M. Automatic Seizure Detection Based on S-Transform and Deep Convolutional Neural Network. *Int. J. Neural Syst.* **2020**, *30*, 1950024.
35. Geng, M.; Zhou, W.; Liu, G.; Li, C.; Zhang, Y. Epileptic Seizure Detection Based on Stockwell Transform and Bidirectional Long Short-Term Memory. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2020**, *28*, 573–580.
36. Kalbkhani, H.; Shayesteh, M.G. Stockwell transform for epileptic seizure detection from EEG signals. *Biomed. Signal Process. Control* **2017**, *38*, 108–118.
37. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. In Proceedings of 31st Annual Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, 4–9 December 2017; pp. 5998–6008.
38. Wang, W.; Xie, E.; Li, X.; Fan, D.-P.; Song, K.; Liang, D.; Lu, T.; Luo, P.; Shao, L. PVT v2: Improved baselines with Pyramid Vision Transformer. *Comput. Vis. Media* **2022**, *8*, 415–424.
39. Wang, Y.Q.; Mohamed, A.; Le, D.; Liu, C.X.; Xiao, A.; Mahadeokar, J.; Huang, H.Z.; Tjandra, A.; Zhang, X.H.; Zhang, F.; et al. Transformer-Based Acoustic Modeling for Hybrid Speech Recognition. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; IEEE: Barcelona, Spain, 2020; pp. 6874–6878.
40. Li, W.; Liu, H.; Ding, R.; Liu, M.; Wang, P.; Yang, W. Exploiting Temporal Contexts With Strided Transformer for 3D Human Pose Estimation. *IEEE Trans. Multimed.* **2023**, *25*, 1282–1293.
41. Sun, J.; Wang, X.; Zhao, K.; Hao, S.Y.; Wang, T.Y. Multi-Channel EEG Emotion Recognition Based on Parallel Transformer and 3D-Convolutional Neural Network. *Mathematics* **2022**, *10*, 3131.
42. Yan, J.Z.; Li, J.N.; Xu, H.X.; Yu, Y.C.; Xu, T.Y. Seizure Prediction Based on Transformer Using Scalp Electroencephalogram. *Appl Sci.* **2022**, *12*, 4158.
43. Li, J.C.; Pan, W.J.; Huang, H.Y.; Pan, J.H.; Wang, F. STGATE: Spatial-temporal graph attention network with a transformer encoder for EEG-based emotion recognition. *Front. Hum. Neurosci.* **2023**, *17*, 1169949.
44. Sun, Y.; Jin, W.; Si, X.; Zhang, X.; Cao, J.; Wang, L.; Yin, S.; Ming, D. Continuous Seizure Detection Based on Transformer and Long-Term iEEG. *IEEE J. Biomed. Health Inform.* **2022**, *26*, 5418–5427.
45. Shoeb, A.H. Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment. Ph.D. Thesis, Massachusetts institute of Technology, Cambridge, MA, USA, 2009.
46. Goldberger, A.L.; Amaral, L.A.N.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.K.; Stanley, H.E. PhysioBank, PhysioToolkit, and PhysioNet—Components of a new research resource for complex physiologic signals. *Circulation* **2000**, *101*, E215–E220.
47. CHB-MIT Scalp EEG Database. Available online: https://physionet.org/content/chbmit/1.0.0/ (accessed on 11 March 2023).
48. Hanley, J.A.; McNeil, B.J. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology* **1982**, *143*, 29–36.
49. Ansari, A.Q.; Sharma, P.; Tripathi, M. Automatic seizure detection using neutrosophic classifier. *Phys. Eng. Sci. Med.* **2020**, *43*, 1019–1028.
50. Janjarasjitt, S. Epileptic seizure classifications of single-channel scalp EEG data using wavelet-based features and SVM. *Med. Biol. Eng. Comput.* **2017**, *55*, 1743–1761.
51. He, J.T.; Cui, J.; Zhang, G.B.; Xue, M.R.; Chu, D.Y.; Zhao, Y.N. Spatial-temporal seizure detection with graph attention network and bi-directional LSTM architecture. *Biomed. Signal Process. Control* **2022**, *78*, 103908.
52. Yao, S.X.; Zhang, Y.L. Transfer Learning and Gated Recurrent Unit Based Epileptic Seizure Detection Method. In Proceedings of the 4th International Conference on Informatics Engineering & Information Science (ICIEIS2021), Tianjin, China, 19–21 November 2021; SPIE: Washington, DC, USA; p. 12161.
53. Cura, O.K.; Akan, A. Classification of Epileptic EEG Signals Using Synchrosqueezing Transform and Machine Learning. *Int. J. Neural Syst.* **2021**, *31*, 2150005.
54. Hu, X.; Yuan, S.; Xu, F.; Leng, Y.; Yuan, K.; Yuan, Q. Scalp EEG classification using deep Bi-LSTM network for seizure detection. *Comput. Biol. Med.* **2020**, *124*, 103919.
55. Duan, L.; Wang, Z.; Qiao, Y.; Wang, Y.; Huang, Z.; Zhang, B. An Automatic Method for Epileptic Seizure Detection Based on Deep Metric Learning. *IEEE J. Biomed. Health Inform.* **2022**, *26*, 2147–2157.
56. Shyu, K.K.; Huang, S.C.; Lee, L.H.; Lee, P.L. Less Parameterization Inception-Based End to End CNN Model for EEG Seizure Detection. *IEEE Access* **2023**, *11*, 49172–49182.
57. Jiang, L.; He, J.; Pan, H.; Wu, D.; Jiang, T.; Liu, J. Seizure detection algorithm based on improved functional brain network structure feature extraction. *Biomed. Signal Process. Control* **2023**, *79*, 104053.
58. Gao, B.; Zhou, J.; Yang, Y.; Chi, J.; Yuan, Q. Generative adversarial network and convolutional neural network-based EEG imbalanced classification model for seizure detection. *Biocybern. Biomed. Eng.* **2022**, *42*, 1–15.

59. Zhang, Y.; Yao, S.; Yang, R.; Liu, X.; Qiu, W.; Han, L.; Zhou, W.; Shang, W. Epileptic Seizure Detection Based on Bidirectional Gated Recurrent Unit Network. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2022**, *30*, 135–145.
60. Yoshiba, T.; Kawamoto, H.; Sankai, Y. Basic study of epileptic seizure detection using a single-channel frontal EEG and a pre-trained ResNet. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2021**, *2021*, 3082–3088.
61. Samiee, K.; Kovács, P.; Gabbouj, M. Epileptic seizure detection in long-term EEG records using sparse rational decomposition and local Gabor binary patterns feature extraction. *Knowl. -Based Syst.* **2017**, *118*, 228–240.
62. Sun, B.; Wu, Z.; Hu, Y.; Li, T. Golden subject is everyone: A subject transfer neural network for motor imagery-based brain computer interfaces. *Neural Netw.* **2022**, *151*, 111–120.
63. Affes, A.; Mdhaffar, A.; Triki, C.; Jmaiel, M.; Freisleben, B. Personalized attention-based EEG channel selection for epileptic seizure prediction. *Expert Syst. Appl.* **2022**, *206*, 117733.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.