

## Research Article

# DWT-Net: Seizure Detection System with Structured EEG Montage and Multiple Feature Extractor in Convolution Neural Network

Zhe Zhang,<sup>1</sup> Yun Ren,<sup>2</sup> Nabil Sabor,<sup>1,3</sup> Jing Pan,<sup>2</sup> Xiaona Luo,<sup>2</sup> Yongfu Li<sup>1</sup>,<sup>1</sup> Yucai Chen<sup>1</sup>,<sup>2</sup> and Guoxing Wang<sup>1</sup>

<sup>1</sup>Department of Micro-Nano Electronics and MoE Key Lab of Artificial Intelligence, Shanghai Jiao Tong University, Shanghai 200240, China

<sup>2</sup>Department of Neurology, Shanghai Jiao Tong University Affiliated Children's Hospital of Shanghai, Shanghai 200062, China

<sup>3</sup>Electrical and Electronics Engineering Department, Faculty of Engineering, Assiut University, Assiut 71516, Egypt

Correspondence should be addressed to Yongfu Li; liyongfu.sg@gmail.com and Yucai Chen; chenyc@shchildren.com.cn

Received 21 November 2019; Revised 3 July 2020; Accepted 30 July 2020; Published 25 August 2020

Academic Editor: Moumen El-Melegy

Copyright © 2020 Zhe Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Automated seizure detection system based on electroencephalograms (EEG) is an interdisciplinary research problem between computer science and neuroscience. Epileptic seizure affects 1% of the worldwide population and can lead to severe long-term harm to safety and life quality. The automation of seizure detection can greatly improve the treatment of patients. In this work, we propose a neural network model to extract features from EEG signals with a method of arranging the dimension of feature extraction inspired by the traditional method of neurologists. A postprocessor is used to improve the output of the classifier. The result of our seizure detection system on the TUSZ dataset reaches a false alarm rate of 12 per 24 hours with a sensitivity of 59%, which approaches the performance of average human detector based on qEEG tools.

## 1. Introduction

Electroencephalograph (EEG) recording refers to the measurement of electrical activity resulting from postsynaptic potentials within the brain [1]. The EEG analysis is used for diagnosing neural dysfunction, such as epileptic seizure, cerebrovascular disease, and brain tumor. With the fast development in data science and machine/deep learning techniques of past decades, automated EEG analysis has great potential in bringing forward an attractive advancement in accuracy and efficiency of diagnosis and treatment of neural dysfunctions [1].

The basic requirement for treatment in different kinds of epilepsy is to identify epileptic seizure features from EEG signals. The equipment for EEG signal acquisition systems is relatively inexpensive; however, the cost for training or hiring a certified neurologist to read and report EEG data is considerably much higher [2]. At the moment, the interpretation of

EEG recordings depends heavily on the judgment of neurologists [3]. It is also time-consuming and tedious to perform 24/7 EEG monitoring to avoid missing epileptic cases [4]. Moreover, the nonstationary nature [5] of EEG and artifacts provides both interrater and intrarater disagreements that degrade diagnosis validity [6]. An automated EEG diagnostic system that provides effective, subjective, and accurate seizure detection is thus needed.

The problem of automatic seizure detection can be divided into two main steps, feature extraction and classifier training. Considerable amounts of works have been done with this two-step procedure for better detection accuracy, including time-frequency feature map with a support vector machine (SVM) [3, 7], nonlinear features with different types of classifiers [5, 8, 9], and features based on time-frequency image with image recognition methods [10, 11]. These researches have provided different methodologies for seizure detection. However, since these works have only been

validated on datasets with less than 25 patients [3], additional verifications are required to validate their methods on a larger dataset. Additional signal processing techniques might be required to fine-tune their methods for practical usage. Recently, an increasing number of researchers start to utilize neural network (NN) methods due to its inherent automatic feature extraction characteristics [12–14]. San-segundo et al. analyzed the use of deep neural network for epileptic EEG signal classification with different inputs and suggested empirical mode decomposition for better performance in focal versus nonfocal classification and Fourier transform for seizure detection [15]. Tsioris et al. presented the long short-term memory (LSTM) method in seizure detection using EEG signals, expanding the use of traditional deep learning algorithms in this field [16]. However, a common problem of these NN methods is that few of them take advantage of existing neurological knowledge to improve the model's accuracy and converge speed.

Controlling of feature extraction provides us with the method to apply neurological knowledge to the model. ChronoNet [17] introduced the concept of multiple-time-scale feature extraction, where inception layers encode time-series information before fusion of the EEG signal. However, this work treats the EEG channels as feature channels at the beginning of the model, failing to learn signal patterns of individual EEG channels. Eberlein et al. [12] performed convolution on EEG signals with kernels ranging over multiple channels to detect local patterns instead of a single channel. Although the authors tried several topologies over the number of channels to be convoluted together, the accuracy is limited due to insufficient representative features in EEG recordings. Both works introduced the idea of manually adjusting the input domain in the early stage of the neural network in seek of better performance of their model. However, to the best of our knowledge, none of the existing research has conducted feature extraction techniques on both temporal and regional patterns as in the field of seizure detection.

To address the aforementioned issues, we propose to use wavelet coefficient packages as input features and introduce the concept of local pattern inception into the neural network model as our seizure detection system. Our model is trained and examined with an up-to-date real clinical EEG dataset [18] that provides a sensitivity of 59.07% at a false alarm rate of 12/24 hours, reaching the average human performance [19].

The key contributions of our work are summarized as follows:

- (1) *Dataset Preparation.* We propose a single reference EEG montage for seizure detection to solve the problem of independence and dependence of adjacent electrodes.
- (2) *Feature Extraction.* We propose a feature extraction method inspired by neurologists' way of reading EEG to improve feature extraction of DWT-Net for seizure detection with efficient computing cost.
- (3) *Neural Network Optimization.* We optimize the kernel size of the convolution layers to deal with differ-

ent temporal resolutions of various discrete wavelet coefficients.

- (4) *Postprocessor.* We optimize the system by concatenating the classifier with a finite-state-machine postprocessor.

The rest of the paper is organized as follows: Section 2 presents preliminaries about EEG recording methods and seizure detection datasets with problem definition. In Section 3, we propose our seizure detection system with detailed methods including feature extraction, network structure, and postprocessor. The experimental results and discussion are presented in Section 4, followed by conclusion in Section 5.

## 2. Preliminaries

In this section, we will discuss the present EEG recording techniques and related datasets and define our problem statement based on the evaluation matrices.

**2.1. EEG Recording Methods.** For clinical epilepsy treatment, noninvasive EEG signal recording is a preferable method because of ethical concerns and medical risks [20]. EEG signals are typically acquired with equipment via the potential difference between pairs of recording electrodes placed on the scalp surface. The measurement between any two electrodes is considered an EGG channel.

According to the “International 10-20 system” measurement standard, the electrodes are distributed across the brain scalp to ensure the reproducibility of EEG experiment [21]. The notation of “10” and “20” defines that the distances between adjacent electrodes are either 10% or 20% of the total longitudinal or transverse distance of the skull. There are a total of 19 recording electrodes and two referential electrodes. Each electrode has a “positional” code and a number; an odd number represents the left brain position, and an even number represents the right brain. As illustrated in Figure 1, the electrode “T3” is termed as temporal lobe (T) on the left side of the brain.

EEG signals are presented in the form of either single reference or bipolar montage, as shown in Figure 2, where a montage is defined as an ordered list of EEG channels recorded in a regular time interval [22]. Both types of montages are used by the clinicians to understand the origin and location of epileptic seizure signals. A single reference montage is also known as a referential montage with one or two referential electrodes. The referential electrodes can be auricular electrodes or averaged potential of all the electrodes. They are paired with the recording electrodes to form the channels in a single reference montage. A bipolar montage does not have referential electrodes, and it records the potential differences between pairs of recording electrodes [23].

**2.2. EEG Datasets.** To develop an automatic seizure detection system, an EEG database with well-defined epileptic recordings is required. Table 1 provides a list of open-source seizure detection datasets (corpora).

The “Bonn” corpus, from the University of Bonn [24], has been widely used for research on seizure detection [8, 11, 25].

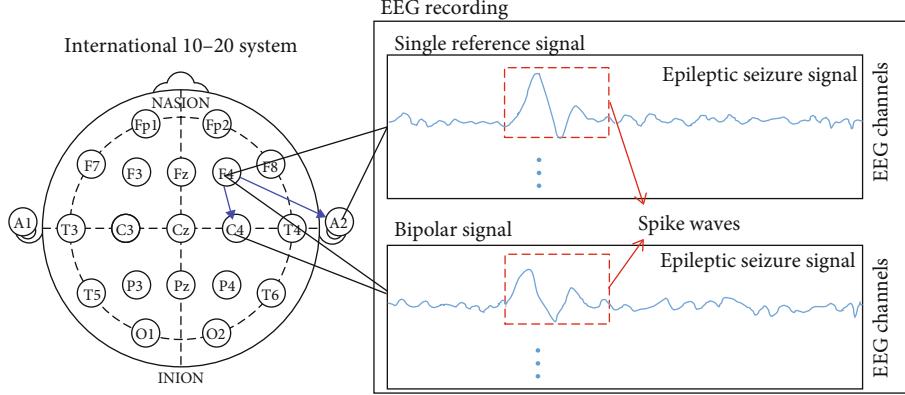


FIGURE 1: Common EEG recording method used for epileptic seizure detection.

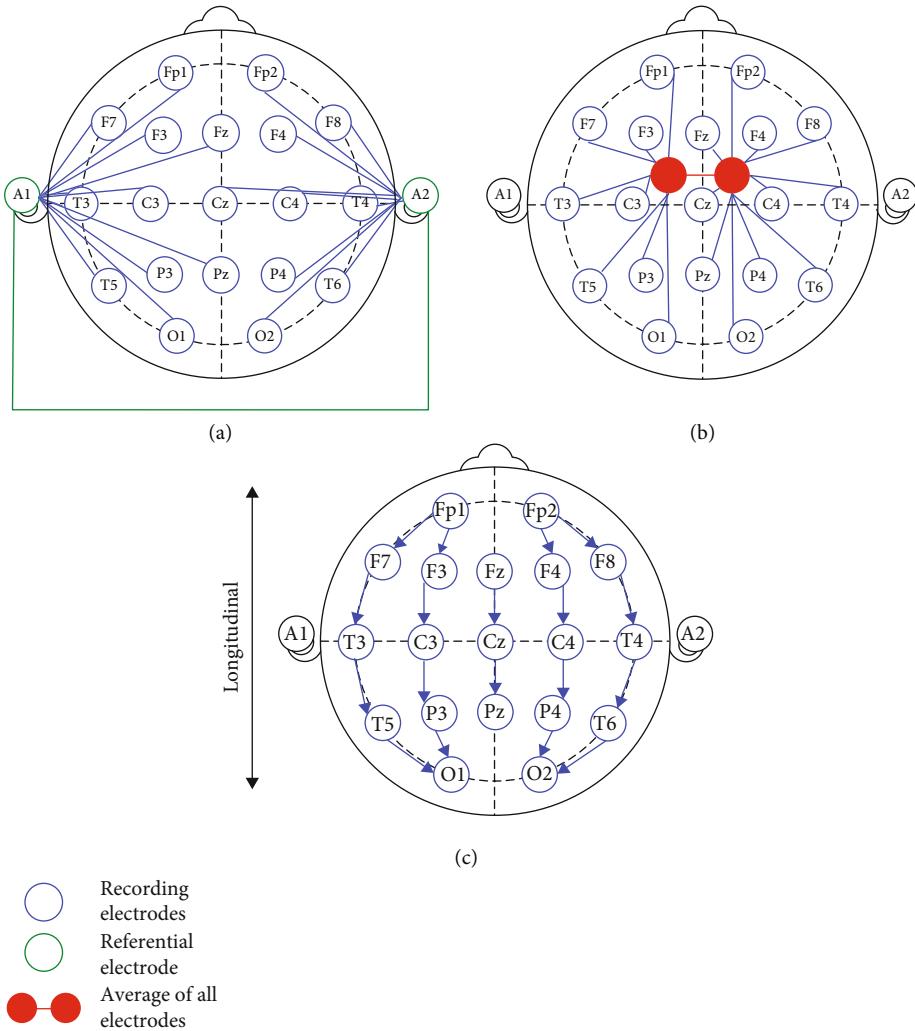


FIGURE 2: Three common EEG montages: (a) the auricular reference montage that uses electrodes on ears as referential electrodes, (b) the common average montage that uses averaged potential of all the electrodes as the referential electrode, and (c) the longitudinal bipolar montage that records the potential difference between pairs of recording electrodes.

It includes single-channel EEG recordings with a total number of 10 subjects and 100 seizure occurrences. All the EEG signals in the corpus were manually reviewed by professional clinicians to pick the representative epileptic channel and to

remove recordings with artifacts. Hence, each session retains only one EEG channel measurement.

The “CHB-MIT” scalp EEG corpus [3, 26] is another widely used dataset for seizure detection [11, 27, 28]. It

TABLE 1: Open source EEG corpus for seizure detection.

	Bonn	CHB-MIT	TUSZ
Number of subjects	10	23	315
Number of seizure occurrence	100	163	1791
Total seizure time	0.6 hours	2.8 hours	36 hours
Total record time	3.2 hours	175 hours	797 hours
% of seizure time/total time	1.88%	1.6%	4.5%
Number of seizure type	1	1	7
Number of channels	1	19	19

consists of continuous scalp EEG recordings from 23 pediatric patients undergoing medication withdrawal for epilepsy surgery evaluation at Children’s Hospital, Boston. The corpus includes 19-channel EEG recordings with a total number of 163 seizure occurrences and a total record time of 175 hours.

The Temple University Hospital Seizure Detection Corpus (TUSZ) [29] is the largest open-source EEG corpus for seizure detection and provides an accurate representation of actual clinical conditions. This corpus is still undergoing updates, and the version used in this paper is v1.5.0. Currently, this corpus includes 19-channel EEG recordings with a total number of 315 subjects, 1791 seizure occurrences, and a total record time of 797 hours. In particular, this corpus is the only dataset that provides different types of epileptic seizure signal. Ref. [30, 31] completed a benchmark on the classification of different types of seizures in TUSZ.

Since “Bonn” and “CHB-MIT” corpuses lack sufficient subjects and data, both corpus might not be a good representative of the real-world clinical situations [32]. Hence, we have adopted the “TUSZ” corpus for our study to develop a seizure detection system.

**2.3. Problem Formulation and Definitions.** We define the following evaluation matrices used to evaluate the performance of a seizure detection system against the “TUSZ” corpus.

**Definition 1. (seizure density function).** An ideal seizure detection system or human marker is expected to label each seizure in the recordings with an accurate start time and end time. An evaluation method [6] is considered to describe the EEG signals by a seizure density function, which varies between 0 and 1 throughout the record. An ideal density function of a detection system is a function of time with the value 1 during the detections and 0 elsewhere.

**Definition 2. (sensitivity (Sen)).** Sensitivity is defined as the ratio between the total number of detected seizure events and the total number of all labelled seizure events in the EEG record,

$$\text{Sen} = \frac{\text{Total number of correctly recognized seizure events}}{\text{Total number of labelled seizure events}}, \quad (1)$$

where a detected seizure event is defined if the average value of seizure density function over the event duration exceeds the threshold  $p$ .

**Definition 3. (specificity (Sp)).** Specificity is defined as the ratio between the number of detected normal event and the total number of all labelled normal events in the EEG record,

$$\text{Sen} = \frac{\text{Total number of recognized normal events}}{\text{Total number of labelled normal events}}, \quad (2)$$

where a normal event is defined if the average value of seizure density function over the event duration falls below the threshold  $p$ .

**Definition 4. (true positive rate (TPR)).** The true positive rate is defined as the ratio between the number of detected seizure events and the total number of events recognized as seizure events,

$$\text{TPR} = \frac{\text{Total number of correctly recognized seizure events}}{\text{Total number of detected seizure events}}. \quad (3)$$

**Definition 5. (false alarm rate (FAR)).** A false alarm occurs if the integral of seizure density function over a nonseizure event exceeds seizure density threshold  $p$ . The false alarm rate over a period of 24 hours is defined as the ratio between the number of false alarm events and the duration of the EEG recorded in the unit of 24 hours.

$$\text{FAR} = \frac{\text{Total number of false alarm events}}{\text{Total duration}}. \quad (4)$$

With the above definitions, the epileptic seizure detection problem is formulated as follows.

**Problem 1. (epileptic seizure detection).** Given a corpus that contains EEG channels with normal and seizure events, train a seizure detection system based on sliding window method. For each sliding window, the sensitivity and specificity of whether it contains a seizure event should be maximized. For the final output of the system, the false alarm rate of seizure events detection should be minimized.

### 3. Feature Extraction Methods and Neural Network Model

#### 3.1. Feature Extraction

**3.1.1. Montage Selection.** The number of channels and the selection of EEG montage have a direct impact on the performance of the classification system [33]. In the TUSZ corpus, it contains more than 40 different channel configurations and 4 different types of reference points. As shown in Figure 3, after our preliminary study on channel selection based on [34, 35], we have decided to use 19 channels from single reference montage based on the “International 10-20 system” to ensure the generality of our model.

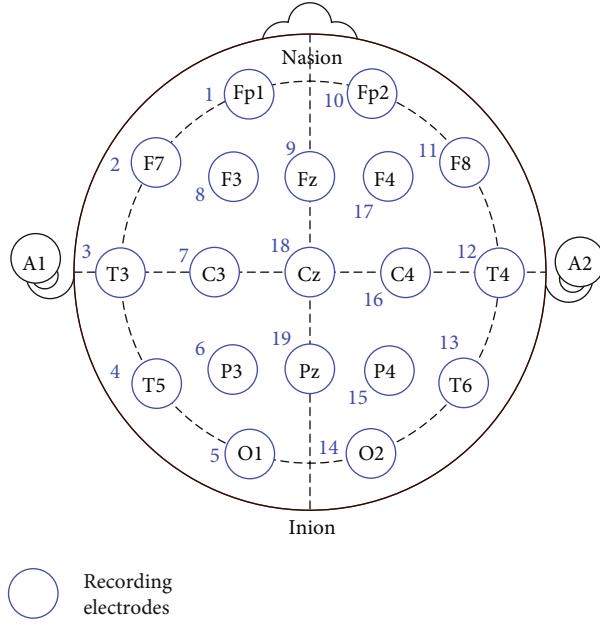


FIGURE 3: Channel sequence of the proposed single reference montage. The sequence is derived from the longitudinal bipolar montage and presented with numbers.

In this work, we have arranged the single reference EEG channels in a new montage to extract the spatial information of the EEG signals. The order of the channels is derived from the longitudinal bipolar montage recommended by standard neurophysiology guidelines [36]. In this montage, each pair of recording electrodes from neighboring channels corresponds to a channel in the longitudinal bipolar montage. For example, the pair of channel Fp1-reference and F7-reference corresponds to Fp1-F7 in the bipolar montage. A bipolar montage can be derived from single reference montages, because the subtraction of two channels with the same referential electrode would cancel the effects of the reference point [37]. Hence, this single reference montage that maintains the bipolar sequence enables the neural network not only to detect features from the single reference electrodes but also make it possible to draw information from the difference between two neighboring electrodes, which mimics the bipolar montage. Compared with other montages, the proposed single reference montage provides the classifier with more information than only using traditional single reference or bipolar montages as the input signal. The related experimental results will be discussed in Section 4.

**3.1.2. Window Length Selection.** To derive the seizure density function from continuous EEG recordings, the EEG signals need to be processed with moving-window analysis. Neurologists typically evaluate the symptom based on 10-second windows of EEG signals [14]. However, a large proportion of the events in the “TUSZ” corpus are shorter than 10 seconds. Thus, we have selected two shorter windows, 1-second and 5-second windows. With the 250 Hz sampling rate of the TUSZ corpus, the 1-second and 5-second windows

contain 19 single reference EEG channels with 250 and 1250 sampling points per channel, respectively.

**3.1.3. EEG Noise Removal and Normalization.** An eighth-order Butterworth filter of 49 to 51 Hz was applied to each window to filter out power-line noise. Each signal was further normalized with Z-score normalization [38], obeying the following equation:

$$X = \frac{x - x_{\text{mean}}}{\sigma_x}, \quad (5)$$

where  $x_{\text{mean}}$  and  $\sigma_x$  represent the mean value and the standard deviation of this EEG within the window duration. The result of Z-score normalization is a signal with zero mean and a standard deviation of 1.

**3.1.4. Discrete Wavelet Transformation.** Discrete wavelet transformation (DWT) is a wavelet decomposition method [39]. This method decomposes a discrete signal into packages of coefficients that represent approximate and detailed information by calculating the inner product of the signal and mother wavelet functions. EEG signals can be seen as a time sequence signal consisting of different frequency components. Thus, the packages of coefficients correspond to the lower and higher frequency component of the signal, respectively. As shown in Figure 4, a detailed package represents a frequency component  $D_{n,m}$ , which corresponds to the frequency range of  $[fs/2^{m+1} \text{ Hz}, fs/2^m \text{ Hz}]$ , where  $fs$  is the sampling rate (250 Hz),  $m = 1, 2, \dots, n$ , and  $n$  is the total level of decomposition.

Most of the related studies use inverse-DWT after DWT to reconstruct EEG into time-sequence signals of different frequency components [5, 25, 28]. In our study, we use the coefficients directly as input features. This method has been tested in other fields such as fault diagnosis [40].

The fourth Daubechies mother wavelet function (db4) is a widely used mother wavelet function in the field of EEG analysis [10, 35, 41]. The morphological characteristic of this mother wavelet function resembles EEG signals. From our preliminary study, we have identified four wavelet coefficient packages decomposed by db4 that represent the lowest frequency components of the EEG, as those frequency components are more related to epileptic signals. The length of the db4 wavelet filter is 8. We can get the length of the decomposed coefficient package, by the following equation

$$\text{Len}(\text{coef } f) = \text{floor} \left[ \frac{(\text{len}(\text{data}) + \text{len}(\text{filter}) - 1)}{2} \right]. \quad (6)$$

The wavelet decomposition stops when the signal becomes shorter than the db4 filter length. Based on the origin signal length 1250 and 250, the maximum decomposition level for 5-second window and 1-second window are 7 and 5, respectively. We take the middle of 1024 sampling points from 1250 sampling points for 5-second windows and 260 sampling points with overlapping of 10 sampling points for 1-second windows. With this method, the length of the coefficient packages representing the lowest frequency components

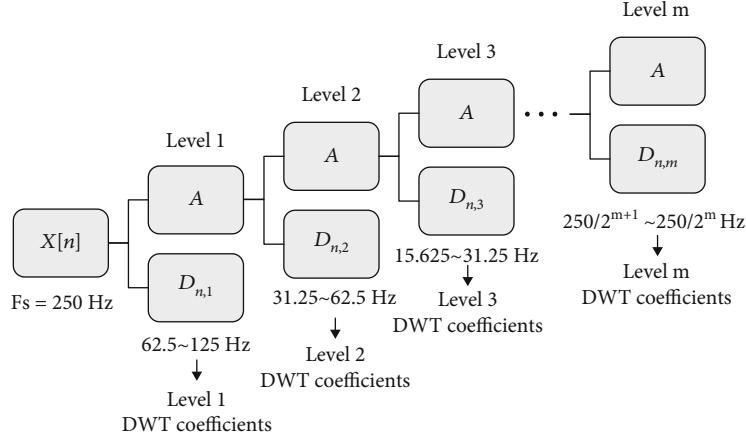


FIGURE 4: An illustration of discrete wavelet transformation used in 5-second window.  $n$  is chosen as 6. For 1-second window, the maximum decomposition level is limited to 5 due to fewer sampling points.

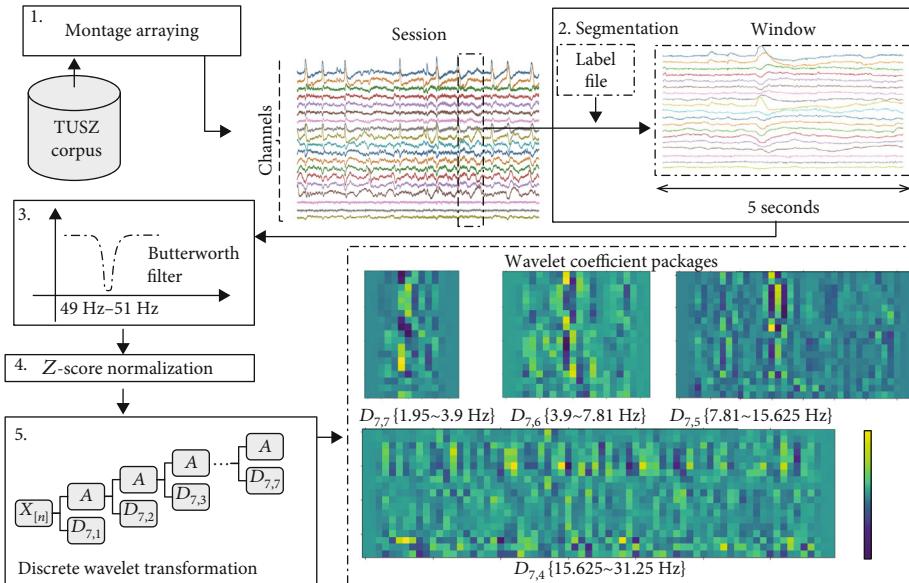


FIGURE 5: Systematic flow of the proposed feature extraction method with 5-second window. Firstly, the EEG signals from TUSZ dataset are restructured into the proposed montage of 19 channels. Secondly, 5-second windows of epileptic and normal EEG are extracted according to the labeling file of the sessions. The windows then go through a Butterworth filter to eliminate power-line interference. After Z-score normalization, the signal is decomposed into discrete wavelet coefficients. Four coefficient packages representing information of lower frequency bands are used as input of the classifier.

for both 1-second and 5-second windows will be 14, 22, 38, and 70, making our model suitable for both window sizes. The proposed feature extraction method is summarized and illustrated in Figure 5.

**3.2. DWT-Net Structure.** Neurologists read EEG signals and recognize abnormal waveforms in the individual channels as well as the correlations between adjacent EEG channels [1]. Their method seeks features from both the time-frequency domain and the spatial domain to ensure that epileptic signals that are unclear in a single domain can be detected.

In this work, we designed a CNN structure called DWT-Net to mimic the above feature extraction methodology. The

model has 9 layers and generates the seizure density of each EEG window. The input of our model is four coefficient packages with sizes of  $C * L_i$  ( $i = 0, 1, 2, 3$ ), where  $C$  represents the number of preprocessed input EEG channels and  $L_i$  represents the length of  $i_{th}$  coefficient package. According to the preprocessed data,  $C$  is 19 and  $L_i$  equals to 14, 22, 38, and 70, respectively.

After the input layer, we implemented multiple feature extractors to process the wavelet packages separately. The method of multiple feature extractors has been used to model local pairwise feature interactions for image recognition [42]. To mimic the abnormal signal identification methodology of neurologists, we extended the concept and implemented a 4-way feature extractor to guide the model. As a result, the

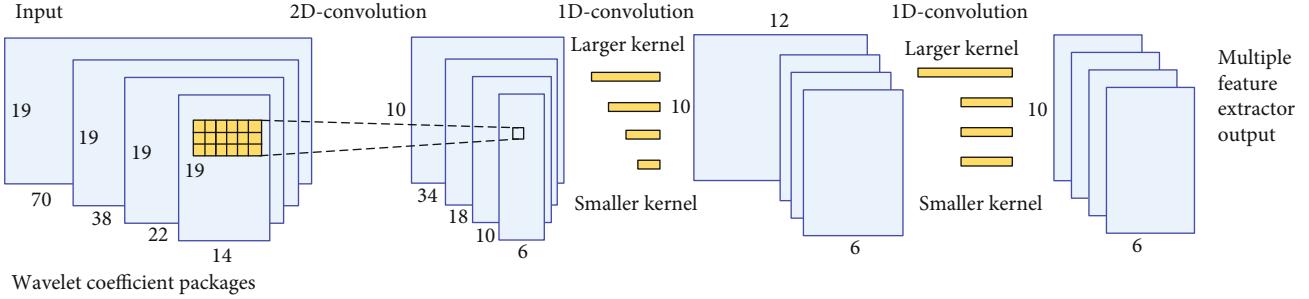


FIGURE 6: Using larger 1D convolution kernel for feature map with larger temporal resolution to normalize the feature maps of the multiple feature extractor.

model will extract features along different domains in a pre-defined order. The second dimension of the feature maps refers to the time steps in the wavelet coefficient packages. It is illustrated in Figure 6 that by adjusting the second dimension of the kernel sizes, strides, and paddings of our convolution layers, we can normalize the temporal resolutions of different DWT coefficient packages. As shown in Figure 7, the details of the model with multiple feature extractors are described as follows:

- (1) 4 wavelet coefficient packages are fed to the 4-input feature extractors
- (2) Each package undergoes the first convolution layer with a kernel size of (3, 6) and stride of 2 for the inception of local temporal and spatial features. The first dimension of the kernel size refers to the number of channels involved in the convolution and is chosen to be 3. With a stride of 2, it ensures that each EEG channel has a chance to interact with neighboring 2 channels. Note that the first convolution layer halves feature dimension along the number of channels
- (3) Two 1d convolution layers are used to extract features along with the temporal dimension of each channel. Different sizes of kernels are used in these two layers between feature extractors for different input coefficient packages. The kernels in size of (1, N) with larger N are used for longer coefficient packages, as shown in Figure 6. The sizes of N, strides, and paddings are carefully chosen to ensure that the shape of output feature maps is normalized correctly
- (4) The output of the third convolution layer is convolved using a kernel size of (3, 3)
- (5) A max-pooling layer is applied to the product of the last kernel, reducing the feature map size to  $5 \times 3$
- (6) The feature map then goes through a dropout layer with a 50% dropout proportion to achieve the effect of auto denoising and prevent overfitting
- (7) The results of 4 feature extractors are stacked by the first dimension into a 3D feature space of  $4 \times 5 \times 3$ .

The first dimension now represents the frequency-domain of the signal

- (8) The last convolution layer with a kernel size of 3 is applied to the 3D feature map. This layer fuse features from different frequency bands together
- (9) Another max-pooling layer is used to reduce the size of the feature map to  $2 \times 3 \times 2$
- (10) The resulting 768 features are fully connected to the output of 2 neurons after two fully connected layers

**3.3. Designing of the Real-Time System.** The softmax output from the classifier is a probability vector of a dimension two, including the probability of the  $n^{\text{th}}$  EEG window to be epileptic EEG  $P_{\text{seizure}}$  and the probability of the  $n^{\text{th}}$  EEG window to be normal EEG  $P_{\text{normal}}$ . As illustrated in Algorithm 1, the postprocessor processes the EEG signal windows by sequence. Starting with state=negative, once the  $n^{\text{th}}$  window is detected as epileptic, the result of the  $n^{\text{th}}$  EEG window Result <sub>$n$</sub>  will be set to Epileptic and the state is changed to Positive (lines 8-9). Under this condition, the system raises the possibility of the next two windows to be epileptic by  $P_{\text{up}}$ , which is set to 0.1 in this work. If the next 2 consecutive windows Result <sub>$n+1$</sub>  and Result <sub>$n+2$</sub>  failed to be detected as Epileptic. The  $n^{\text{th}}$  window will be regarded as a false alarm and Result <sub>$n$</sub>  will be revised to Normal (line 25). Likewise, a single Normal window between two Epileptic windows will be revised to Epileptic. Although such postprocessor brings about a latency of 10 seconds, the proposed postprocessor smooths the sequential hypotheses of the classifiers and provides a moderate effect in suppressing FAR and increasing Sen.

## 4. Experiment Results and Discussion

**4.1. Dataset Preparation.** In this work, we have chosen the latest version v1.5.0 of the “TUSZ” corpus. In TUSZ, the EEG labels are given with a start time and an end time for each EEG recording. The resolution of the labels in the TUSZ dataset is 0.0001 second. We can define an event with its start time and end time. The sessions of the corpus that includes eight classes of events (“background” event and 7 “epileptic” events) are listed in Table 2. As we are not going to include seizure classification in our system, all the epileptic classes

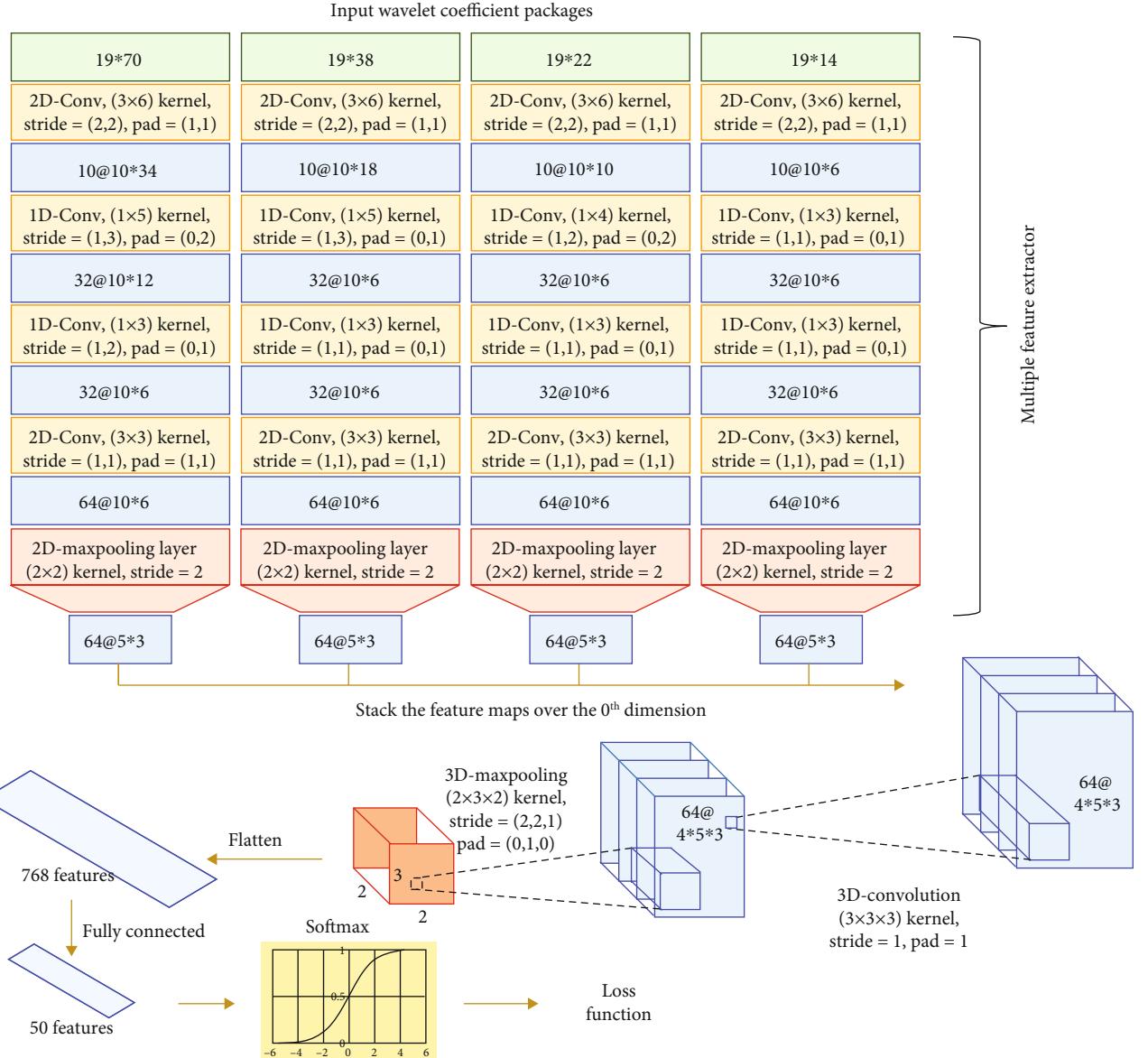


FIGURE 7: Structure of DWT-Net. The upper part illustrates the structure of multiple feature extractor proposed for better feature extraction performance with regard to epileptic EEG signal. In the bottom part, outputs of the multiple feature extractor are stacked into a 3D feature map, which goes through common convolution layers for classification.

are considered “positive” events and the background class “negative” events in the experiment, respectively. Because the window length of our seizure detection system is only 5 seconds, we regard seizure events with short breaks as separate events to make the best use of the labels.

The first minute of an EEG recording provides tail-tail signs whether the signal is epileptic or not because the epileptic signal is strongest at the beginning of a seizure [17]. Hence, limiting the length of the events can provide better classification performance. However, extracting only 60 seconds of signal from the events will result in shortage of training data required for training a deep learning model and make the classifier less adaptable. To balance the amount of training data and classifier performance, for events greater than 400 seconds, only the first 400 seconds are included.

A total number of 6971 and 2238 sessions were used as the train and test sets predefined by TUSZ. After feature extraction, 119,491 5-second windows were generated with details showed in Table 3. Correspondingly, the 1-second windows were derived from the 5-second windows.

**4.2. Training of Window-Based Classifier.** The proposed model was constructed with the open-source framework Pytorch [43]. The weights of the neural network were initialized with Kaiming normalization [44] to improve weight convergence during the training of our model with ReLu activation layers. The weighted cross-entropy was selected as the loss function of the classifier. We have used the Adam optimization method [45] with  $\beta_1$  of 0.9,  $\beta_2$  of 0.999, learning rate of 0.0005, and weight decay (L2 penalty regularization)

```

Require: Pseizure, Pnormal, Pup
Output: Resultn, n=1,2,...
1: Initialize System State State  $\leftarrow$  Negative
2: while New EEG window do
3:   if State == Negative then
4:     if Pseizure < Pnormal then
5:       next State  $\leftarrow$  Negative
6:       Resultn  $\leftarrow$  Normal
7:     else
8:       next State  $\leftarrow$  Positive
9:       Resultn  $\leftarrow$  Epileptic
10:    end if
11:   end if
12:   if State == Positive then
13:     if Pseizure + Pup < Pnormal then
14:       next State  $\leftarrow$  Smooth
15:       Resultn  $\leftarrow$  Normal
16:     else
17:       next State  $\leftarrow$  Positive
18:       Resultn  $\leftarrow$  Epileptic
19:     end if
20:   end if
21:   if State == Smooth then
22:     if Pseizure + Pup < Pnormal then
23:       next State  $\leftarrow$  Negative
24:       Resultn  $\leftarrow$  Normal
25:       Resultn-2  $\leftarrow$  Normal
26:     else
27:       next State  $\leftarrow$  Positive
28:       Resultn  $\leftarrow$  Positive
29:       Resultn-1  $\leftarrow$  Positive
30:     end if
31:   end if
32: end while

```

ALGORITHM 1. Algorithm for state machine post-processor

TABLE 2: Number of sessions for classes in TUSZ v1.5.0.

Class	Train set	Test set
Background	5313	1597
Focal nonspecific seizure	1015	283
Generalized nonspecific seizure	342	175
Complex partial seizure	147	59
Tonic seizure	18	51
Simple partial seizure	49	3
Tonic clonic seizure	37	21
Absence seizure	50	49

of 0.005. The Adam method combines the advantages of AdaGrad [46] and RMSprop [47], and it automatically adjusts the learning rate during training to accelerate the convergence of the model. The learning rate is adjusted according to TPR result between the default value 0.001 [43] and 0.0005, which is a similar technique reported in [14]. Similarly, the weight decay is adjusted from the default value 0 to 0.005 which is a similar technique reported in [17] to achieve better performance. We use default settings given

TABLE 3: Number of 5-second windows in the experiment.

	Negative	Positive
Train set	71674	15542
Test set	26058	6217

TABLE 4: Classifier performance on TUSZ.

Classifier	Sensitivity	Specificity	TPR
5 sec window origin dataset	45.69%	76.06%	23.11%
1 sec window origin dataset	35.86%	83.25%	28.59%
5 sec window 5-fold	77.23%	83.41%	37.78%
1 sec window 5-fold	79.10%	81.38%	34.61%
5 sec window with +1 conv layers	76.28%	80.91%	34.90%
5 sec window with +2 conv layers	77.82%	81.33%	34.35%
5 sec window without SMOTE	67.89%	82.45%	33.49%
5 sec window double banana montage	70.98%	73.17%	33.34%

by Pytorch for other parameters (e.g.,  $\beta_1$  and  $\beta_2$ ) as these parameters have a negligible effect on the learning speed or the accuracy based on our experimental trials.

The synthetic minority oversampling technique (SMOTE) [48] is adopted to deal with the imbalanced dataset problem between positive and negative classes. SMOTE generates samples in a minority class by calculating an interpolation between a randomly selected minority class sample and one of its  $k$ -nearest minority class neighbors. With a batch size of 64, 56 samples were taken from the dataset and eight minor class samples were generated. This method decreases the ratio between negative and positive samples from 2.25 to 1.6.

Each classifier in our experiment undergoes 180 epochs during the training. The performance comparison between different classifiers is shown in Table 4. TUSZ provides an original segmentation of the training and validation datasets. We shuffled the dataset to create a random dataset for  $k$ -fold evaluation. The results based on the original dataset and  $k$ -fold evaluation were both examined, and  $k$  was chosen to be 5 so that the number of samples in the training phase and the validation phase is similar to the original dataset. The result of the  $k$ -fold evaluation is remarkably better than the result of the original dataset on Sen and TPR. This phenomenon can be explained through the  $k$ -fold dataset is shuffled to balance the proportion of different classes of epileptic events in the training and validation sets, which improve the intrinsic performance of the dataset. Thus, the rest of the results discussed in the paper were all obtained with  $k$ -fold evaluation.

The best overall result was obtained by the 5-second window classifier with the SMOTE technique, whose performance versus epoch is shown in Figure 8. The sensitivity of the 5-second window classifier without SMOTE will drop by 9% with similar specificity performance. Deeper models based on DWT-Net were constructed by increasing the number of convolution layers by 1 to 2 in the network after the

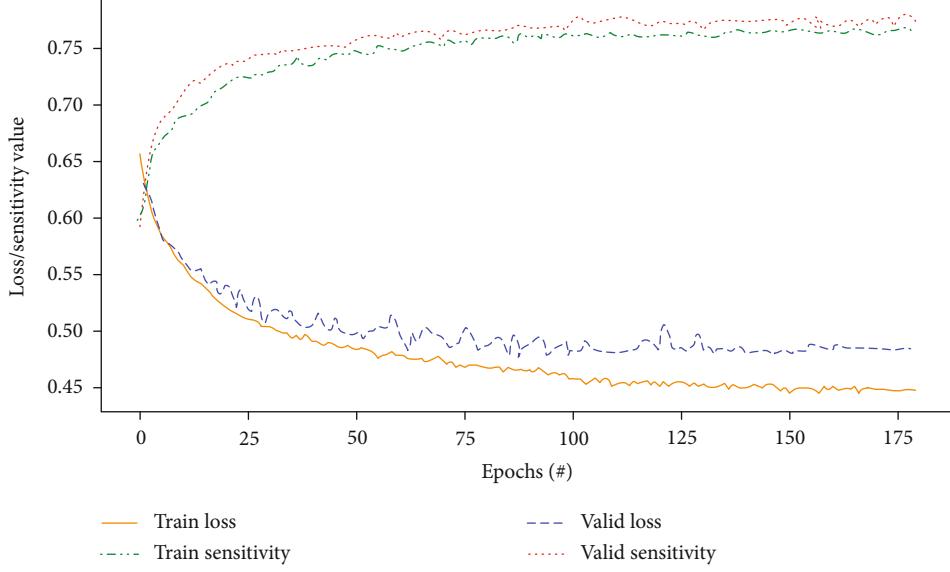


FIGURE 8: The loss and sensitivity values during training and validation of our proposed 5-second-window DWT-Net classifier.

TABLE 5: Results of the proposed seizure detection system in comparison with results reported in other work.

	System/ $p$ value	Sensitivity	Specificity	FAR
Deep learning methods in [14]	HMM	30.32%	80.07%	244
	HMM/SdA	35.35%	73.35%	77
	HMM/LSTM	30.05%	80.53%	60
	IPCA/LSTM	32.97%	77.57%	73
	CNN/MLP	39.09%	76.84%	77
	CNN/LSTM	30.83%	96.86%	7
	$p = 80\%$	25.44%	97.05%	3
Our DWT-Net system based on 5-second DWT-Net	$p = 76\%$	30.25%	96.64%	4
	$p = 70\%$	33.45%	96.10%	5
	$p = 60\%$	42.35%	94.93%	6
	$p = 35\%$	59.07%	89.72%	12
	$p = 90\%$	19.40%	95.01%	6
Our DWT-Net system based on 1-second DWT-Net	$p = 83.2\%$	30.07%	93.83%	7
	$p = 80\%$	33.45%	96.10%	5
	$p = 70\%$	42.35%	94.93%	6
	$p = 62.6\%$	49.29%	72%	12

inception layers of the multiple feature extractor. According to the experimental results shown in Table 4, the additional convolution layers cannot provide any improvement in the classifier's performance. It can be concluded that our proposed model can effectively learn features from the input and thus reduce the complicity of a neural network model.

The classifier's ability to learn spatial information from our proposed montage is tested by another trial using the double banana montage. The TPR result without our proposed montage drops by 4.4%. Hence, with the combined techniques (our proposed montage and the classifier), we can improve the Sen of the classifier by 16%.

**4.3. Evaluation of Seizure Detection System.** An epileptic event is defined when the average seizure density function value over its duration is higher than the threshold  $p$ . This output is compared with labeled ground truth for the calculation of Sen and FAR. The results based on 1-second and 5-second detection windows are shown in Table 5. Different values of the threshold  $p$  were examined to characterize the system. Given the same  $p$ , the performance of the system based on the 5-second window classifier performs better than 1-second system by an average increase of 7% in sensitivity. The specificity of the 1-second system drops sharply when the threshold decreases, while the 5-second system still holds

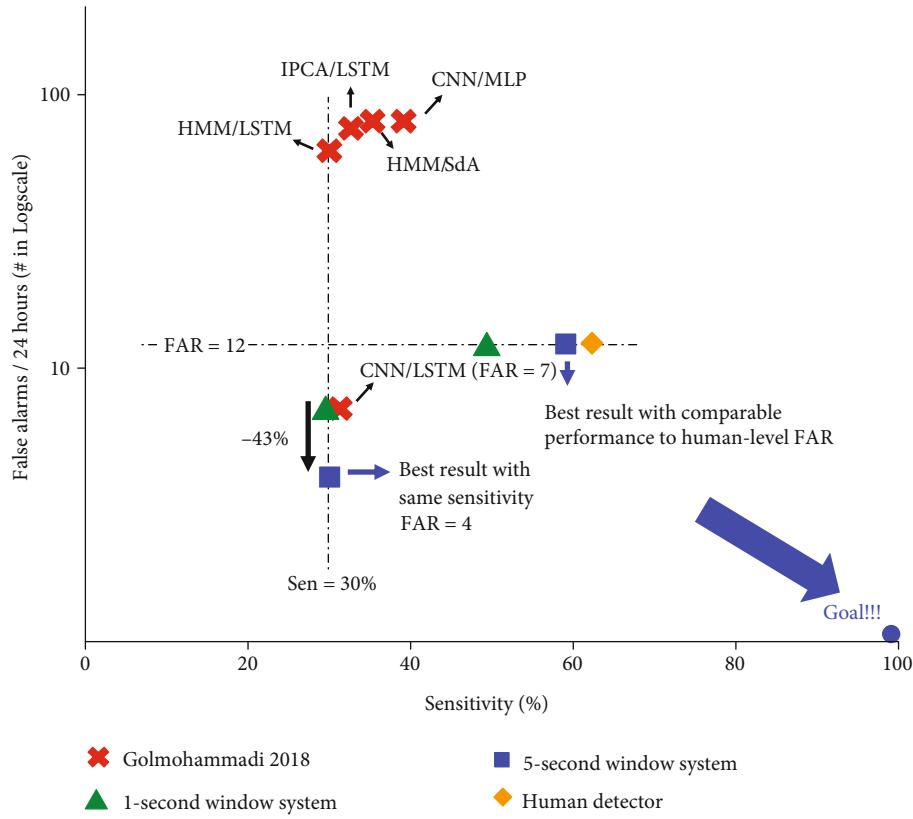


FIGURE 9: FAR versus Sen. Our system based on 5-second DWT-Net classifier achieves the best FAR of 4/24 hours with the sensitivity level of 30%, exceeding the CNN/LSTM model by a 43% decrease in FAR. The method also approaches average human-level detector based on the qEEG tool with a Sen of 59.07% at FAR of 12/24 hours.

a specificity of 89.72% with a 0.35 threshold. We conclude that our model can obtain better features from a 5-second window classifier, whose input wavelet coefficients represent more information at the lower frequency band of 2 to 32 Hz. This result agrees with clinical knowledge that the representative frequency band of seizure is below 30 Hz, proving the model's ability to learn meaningful representative features of EEG signals with the proposed method.

Comparing with the state-of-the-art works [14], we have provided a figure of FAR versus sensitivity in Figure 9 to illustrate the advantage of our proposed system. The system based on 5-second DWT-Net classifier achieves the best FAR of 4/24 hours with the same sensitivity level of 30.25%, exceeding the CNN/LSTM model by 43% decrease in FAR. It is reported by [19] that for similar tasks, the average human performance based on qEEG tools is within the range of 65% sensitivity with a FAR of 12 per 24 hours. To compare the system with human performance, the FAR is fixed to 12/24 hours by adjusting the density threshold  $p$ , and our system can detect seizure events with a sensitivity of 59.07%. The proposed system is almost reaching the performance of an expert clinician with a diagnostic tool.

## 5. Conclusion

In this work, the cooperative design of a multiple feature extractor CNN structure with wavelet coefficient packages

as input is derived from the proposed EEG montage. We introduced a system for automatic real-time detection of epileptic EEG events. Multiple feature extractors are used in our proposed DWT-Net to guide the feature extraction behavior of the model and improve its ability to inception local temporal and spatial features. A sensitivity of 59% is obtained with a FAR of 12/24 hours. While the system improves the state-of-the-art result of the automatic seizure detector to nearly a human level, it does not require an additional computation cost or more neural network layers. Our proposed method achieves similar performance compared to the human-level detector with qEEG tools. In practice, however, the combination of qEEG tools and raw EEG used by neurologists provides better accuracy. Our system can be more robust with additional verification done by the neurologists.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that we have no financial or personal relationship with other people or organizations that could inappropriately influence or bias the content of this paper.

## Acknowledgments

This research is supported in part by the National Key Research and Development Program of China under Grant No. 2019YFB2204500 and in part by the National Science Foundation of China under Grant No. 61874171 and in part by the Science, Technology and Innovation Action Plan of Shanghai Municipality, China, under Grant No. 1914220370.

## References

- [1] D. L. Schomer and F. H. Lopes da Silva, *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, Oxford University Press, Oxford, UK, 2017.
- [2] S. López, G. Suarez, D. Jungreis, I. Obeid, and J. Picone, "Automated identification of abnormal adult EEGs," in *2015 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, Philadelphia, PA, USA, December 2015.
- [3] A. H. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," PhD thesis, Massachusetts Institute of Technology, 2010.
- [4] M. Sharma, R. B. Pachori, and U. Rajendra Acharya, "A new approach to characterize epileptic seizures using analytic time-frequency flexible wavelet transform and fractal dimension," *Pattern Recognition Letters*, vol. 94, pp. 172–179, 2017.
- [5] U. R. Acharya, S. V. Sree, P. C. A. Ang, R. Yanti, and J. S. Suri, "Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals," *International Journal of Neural Systems*, vol. 22, no. 2, article 1250002, 2012.
- [6] S. B. Wilson, M. L. Scheuer, C. Plummer, B. Young, and S. Pacia, "Seizure detection: correlation of human experts," *Clinical Neurophysiology*, vol. 114, no. 11, pp. 2156–2164, 2003.
- [7] A. Rizal and S. Hadiyoso, "Sample entropy on multidistance signal level difference for epileptic EEG classification," *The Scientific World Journal*, vol. 2018, Article ID 8463256, 6 pages, 2018.
- [8] S. Ghosh-Dastidar and H. Adeli, "Improved spiking neural networks for EEG classification and epilepsy and seizure detection," *Integrated Computer-Aided Engineering*, vol. 14, no. 3, pp. 187–212, 2007.
- [9] M. D. K. Hasan, M. A. Ahmed, M. Ahmad, and M. A. Rashid, "Prediction of Epileptic Seizure by Analysing Time Series EEG Signal Using -NN Classifier," *Applied Bionics and Biomechanics*, vol. 2017, Article ID 6848014, 12 pages, 2017.
- [10] Y. Cao, Y. Guo, H. Yu, and X. Yu, "Epileptic seizure auto-detection using deep learning method," in *2017 4th International Conference on Systems and Informatics (ICSAI)*, Hangzhou, China, November 2017.
- [11] Y. Li, W. Cui, M. Luo, K. Li, and L. Wang, "Epileptic seizure detection based on time-frequency images of EEG signals using Gaussian mixture model and gray level co-occurrence matrix features," *International Journal of Neural Systems*, vol. 28, no. 7, article 1850003, 2018.
- [12] M. Eberlein, R. Hildebrand, R. Tetzlaff et al., "Convolutional neural networks for epileptic seizure prediction," in *IEEE International Conference on Bioinformatics and Biomedicine*, Madrid, Spain, December 2018.
- [13] M. Golmohammadi, S. Ziyabari, V. Shah et al., "Gated recurrent networks for seizure detection," in *Signal Processing in Medicine and Biology Symposium*, Philadelphia, PA, USA, December 2017.
- [14] M. Golmohammadi, S. Ziyabari, V. Shah, S. L. de Diego, I. Obeid, and J. Picone, "Deep architectures for automated seizure detection in scalp EEGs," in *International Conference on Machine Learning*, Stockholm, Sweden, July 2017.
- [15] R. San-Segundo, M. Gil-Martín, L. F. D'Haro-Enríquez, and J. M. Pardo, "Classification of epileptic EEG recordings using signal transforms and convolutional neural networks," *Computers in Biology and Medicine*, vol. 109, no. 6, pp. 148–158, 2019.
- [16] K. M. Tsioris, V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis, "A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals," *Computers in Biology and Medicine*, vol. 99, pp. 24–37, 2018.
- [17] S. Roy, I. Kiral-Kornek, and S. Harrer, "ChronoNet: a deep recurrent neural network for abnormal EEG identification," *Artificial Intelligence in Medicine*, vol. 11526, pp. 47–56, 2019.
- [18] V. Shah, E. von Weltin, S. Lopez et al., "The temple university hospital seizure detection corpus," *Frontiers in Neuroinformatics*, vol. 12, p. 83, 2018.
- [19] C. B. Swisher, C. R. White, B. E. Mace et al., "Diagnostic accuracy of electrographic seizure detection by neurophysiologists and non-neurophysiologists in the adult ICU using a panel of quantitative EEG trends," *Journal of Clinical Neurophysiology*, vol. 32, no. 4, pp. 324–330, 2015.
- [20] F. Carpi and D. De Rossi, "Non invasive brain-machine interfaces," *Ariadna Final Report*, vol. 31, pp. 1–30, 2006.
- [21] R. W. Homan, J. Herman, and P. Purdy, "Localisation cérébrale des électrodes placées selon le système international 10-20," *Electroen- cephalography and Clinical Neurophysiology*, vol. 66, no. 4, pp. 376–382, 1987.
- [22] "How do reference montage and electrodes setup affect the measured scalp EEG potentials?," *Journal of Neural Engineering*, vol. 15, no. 2, article 026013, 2018.
- [23] M. Sazgar and M. G. Young, "Overview of EEG, electrode placement, and montages," in *Absolute Epilepsy and EEG Rotation Review*, pp. 117–125, Springer International Publishing, 2019.
- [24] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state," *Physical Review E*, vol. 64, no. 6, article 061907, 2001.
- [25] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 9, pp. 1545–1551, 2007.
- [26] A. L. Goldberger, L. A. Amaral, L. Glass et al., "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. E215–E220, 2000.
- [27] P. Thodoroff, J. Pineau, and A. Lim, "Learning robust features using deep learning for automatic seizure detection," in *Proceedings of the 1st Machine Learning for Healthcare Conference*, Los Angeles, CA, USA, August 2016.
- [28] A. Bhattacharyya and R. B. Pachori, "A multivariate approach for patient specific EEG seizure detection using empirical wavelet transform," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 9, pp. 2003–2015, 2017.

- [29] A. Harati, S. López, I. Obeid, J. Picone, M. Jacobson, and S. Tobochnik, "The TUH EEG CORPUS: a big data resource for automated EEG interpretation," in *2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, Philadelphia, PA, USA, December 2014.
- [30] I. R. D. Saputro, N. D. Maryati, S. R. Solihati, I. Wijayanto, S. Hadiyoso, and R. Patmasari, "Seizure type classification on EEG signal using support vector machine," *Journal of Physics: Conference Series*, vol. 1201, article 012065, 2019.
- [31] S. Roy, U. Asif, J. Tang, and S. Harrer, "Machine learning for seizure type classification: setting the benchmark," 2019, <https://arxiv.org/abs/1902.01012>.
- [32] M. Golmohammadi, A. H. Harati Nejad Torbati, S. Lopez de Diego, I. Obeid, and J. Picone, "Automatic analysis of EEGs using big data and hybrid deep learning architectures," *Frontiers in Human Neuroscience*, vol. 13, no. 3, p. 76, 2019.
- [33] S. López, A. Gross, S. Yang, M. Golmohammadi, I. Obeid, and J. Picone, "An analysis of two common reference points for EEGs," in *2016 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, Philadelphia, PA, USA, December 2016.
- [34] J. Duun-Henriksen, T. W. Kjaer, R. E. Madsen, L. S. Remvig, C. E. Thomsen, and H. B. D. Sorensen, "Channel selection for automatic seizure detection," *Clinical Neurophysiology*, vol. 123, no. 1, pp. 84–92, 2012.
- [35] M. Qaraqe, M. Ismail, and E. Serpedin, "Band-sensitive seizure onset detection via CSP-enhanced EEG features," *Epilepsy & Behavior: E&B*, vol. 50, no. 7, pp. 77–87, 2015.
- [36] J. N. Acharya, A. J. Hani, P. Thirumala, and T. N. Tsuchida, "American clinical neurophysiology society guideline 3: a proposal for standard montages to be used in clinical EEG," *The Neurodiagnostic Journal*, vol. 56, no. 4, pp. 253–260, 2017.
- [37] D. Hagemann, E. Naumann, and J. F. Thayer, "The quest for the EEG reference revisited: a glance from brain asymmetry research," *Psychophysiology*, vol. 38, no. 5, pp. 847–857, 2010.
- [38] U. R. Acharya, H. Fujita, S. L. Oh et al., "Deep convolutional neural network for the automated diagnosis of congestive heart failure using ecg signals," *Applied Intelligence*, vol. 49, no. 1, pp. 16–27, 2019.
- [39] A. Teolis, *Discrete wavelet transform*, Birkhäuser, Boston, 1998.
- [40] M. Zhao, M. Kang, B. Tang, and M. Pecht, "Multiple wavelet coefficients fusion in deep residual networks for fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 6, pp. 4696–4706, 2019.
- [41] N. Rafiuddin, Y. Khan, and O. Farooq, "Feature extraction and classification of EEG for automatic seizure detection," in *2011 International Conference on Multimedia, Signal Processing and Communication Technologies*, Aligarh, India, December 2011.
- [42] T.-Y. Lin, A. R. Chowdhury, and S. Maji, "Bilinear CNN models for fine-grained visual recognition," in *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, December 2015.
- [43] A. Paszke, S. Gross, S. Chintala et al., "Automatic differentiation in PyTorch," in *IEEE Conference on Neural Information Processing Systems*, Long Beach, USA, December 2017.
- [44] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on ImageNet classification," in *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, December 2015.
- [45] D. Kingma and J. Ba, "Adam: a method for stochastic optimization," in *International Conference on Learning Representations*, San Diego, CA, USA, May 2015.
- [46] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," *Journal of Machine Learning Research*, vol. 12, no. 7, pp. 257–269, 2011.
- [47] L. Bottou and Y. LeCun, "Large scale online learning," in *Advances in Neural Information Processing Systems 16 (NIPS 2003)*, pp. 217–225, MIT Press, Cambridge, MA, 2004.
- [48] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.