

Research Article

Automatic Seizure Detection Using Multi-Input Deep Feature Learning Networks for EEG Signals

Qi Sun , Yuanjian Liu , and Shuangde Li 

College of Electronic and Optical Engineering and College of Flexible Electronics (Future Technology),
Nanjing University of Posts and Telecommunications, Nanjing, China

Correspondence should be addressed to Yuanjian Liu; liuyj@njupt.edu.cn and Shuangde Li; lishuangde@njupt.edu.cn

Received 19 June 2023; Revised 29 December 2023; Accepted 22 January 2024; Published 5 February 2024

Academic Editor: Lihang Feng

Copyright © 2024 Qi Sun 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.

Epilepsy, a neurological disease associated with seizures, affects the normal behavior of human beings. The unpredictability of epileptic seizures has caused great obstacles to the treatment of the disease. The automatic seizure detection method based on electroencephalogram (EEG) can assist experts in predicting seizures to improve treatment efficiency. Epileptic seizure detection cannot be achieved accurately using the single-view characteristics of the signals. Moreover, manual feature extraction is a time-consuming task. To design a high-performance seizure identification method, automatic learning of multi-view features becomes an indispensable part for seizure detection. Therefore, the paper proposes a multi-input deep feature learning networks (MDFLN) model, which comprehensively considers the features from the time domain and the time–frequency (TF) domain for EEG signals. The MDFLN model automatically extracts the feature information of the signals through deep learning networks. Then, the bidirectional long short-term memory (BLSTM) network is used to distinguish seizure and nonseizure events. Furthermore, the effectiveness of the proposed network structure is verified in two public datasets. The experimental results demonstrate that the classification accuracy of the proposed method based on multi-view features is at least 2.2% higher than the single-view features. The MDFLN achieves better performance on CHB-MIT and Bonn datasets with accuracy of 98.09% and 98.4%, respectively. The fine-tuned model with the validation set also improves the classification performance. Compare with the state-of-the-art seizure detection methods, the multi-input deep learning network has superior competence with high sensitivity on the CHB-MIT dataset. The proposed automatic seizure detection method can reduce time consumption and effectively assist experts in the clinical diagnosis and treatment.

1. Introduction

Epilepsy is a neurological disease characterized by a sudden rush of electrophysiological signals changing inside the brain [1–3]. More than 60 million people worldwide suffer from different types of epilepsy, especially in developing countries [4]. Seizure detection is an important task in clinical research, which motivates a great deal of research on developing and testing automatic seizure detection algorithms to make clinical strategies [5]. In addition, predicting the onset of seizures can help these patients with further treatment [6, 7]. Scalp electroencephalogram (EEG) is an important tool for the diagnosis of patients with epilepsy [8]. In recent years, digital EEG monitoring systems have collected long-term EEG recordings of epileptic patients in real time to identify the occurrence of abnormal events and make decisions on time.

Experts identify epileptic events by reading long-term EEGs, which is a time-consuming task. Automated seizure monitoring technology can help experts to identify epileptic events in EEG signals.

So far, many works on seizure detection and prediction are based on the traditional machine learning methods. In machine learning, researchers consider using EEG data to extract the features of the EEG signals from the time domain, frequency domain, time–frequency (TF) domain, and non-linear domain [9–11]. These features are used as input to a classifier to detect and classify EEG signals. Commonly used machine learning algorithms include support vector machine (SVM), random forest (RF), K-nearest neighbors (K-NN), and artificial neural network (ANN). The TF method combines time–frequency information to analyze the TF distribution of time series signals. Considering the multi-channel

and instability characteristics of EEG signals, the short-time Fourier transform (STFT) or the empirical wavelet transform (EWT) is often used to analyze the signals. Tzallas et al. [12, 13] used the TF analysis method to classify EEG signals. They used STFT and some TF distributions to calculate the power spectrum density (PSD) of the signal. Bhattacharyya and Pachori [14] explored EWT and designed a TF plane for EEG signals by using joint instantaneous amplitudes and frequencies function in adaptive frequency scale of signal. Chowdhury et al. [15] used another signal decomposition method, named empirical mode decomposition (EMD), to transform the time-domain signals into several amplitude and frequency signals. The bimodal Gaussian model was used to extract the signal information. There are several methods for processing EEG based on the decomposition and reconstruction of signals. Zabini et al. [16] used the time-delay embedding method to reconstruct the trajectories of seizure and nonseizure signals in a high-dimensional space for deep analysis. Jiang et al. [17] used a symplectic geometric method to obtain simplified eigenvalues, which were regarded as input of SVM for the classification of EEG signals. Feature optimization algorithms were also used to improve the performance of the seizure detection method. Subasi et al. [18] proposed a hybrid machine learning method using a genetic algorithm and particle swarm optimization to find the optimal parameters of SVMs. Our previous work focused on finding an optimal feature set for the machine learning algorithm to reduce computational time for seizure detection [19]. However, extracting these features requires manual processing, and one cannot perform deep learning on these features. Most of these works require feature extraction techniques before using machine learning classifiers for seizure detection, which is a major shortcoming of machine learning. In this paper, a method for automatic seizure detection using deep learning networks is developed.

Deep learning networks have satisfied performance for EEG based seizure detection in the presence of noise. During the collection of EEG signals, the presence of artificial noise also makes seizure detection a great challenge. To overcome these difficulties, deep learning technology has emerged, which can automatically learn the relevant features of EEG signals without feature engineering through a supervised learning framework. Many existing studies have proved the effectiveness of deep learning in the classification of EEG signals [20–24]. So far, deep learning has developed rapidly in visual recognition. However, due to the nonstationary nature of EEG signals, the performance of deep learning in seizure detection still needs to be improved. Noise in EEG signals causes decomposed representations to be commonly used as input to the deep learning algorithms. Decomposition is usually performed in the form of a Fourier or wavelet transform. Deep learning algorithms are considered to be easier to obtain relevant features from these decompositions.

Convolutional neural network (CNN) automatically learn features from EEG signals without the need for manual feature extraction. Machine learning techniques transform data into continuous representation spaces, usually using

simple transformations such as high-dimensional nonlinear maps or decision trees. But these techniques often fail to obtain an accurate representation of the complex problems. Therefore, machine learning must make the initial data more suitable for these methods and manually design the presentation layer for the data, which is called feature engineering. In contrast, deep learning can learn all the features without manual design. This greatly simplifies the machine learning workflow and replaces a complex multi-stage process with a simple, end-to-end deep learning model. Truong et al. [21] used STFT to convert the segmented signal into a TF image, and CNN to automatically learn the feature information of the video image to better classify preictal and interictal signals. Tian et al. [22] designed an interpretable rule-based model with multi-view fuzzy system. The method first used fast Fourier transform (FFT) and wavelet packet decomposition (WPD) to obtain the multi-view features which were then inputted to CNN for deep feature learning [22]. O’Shea et al. [23] designed a fully coherent network architecture for neonatal seizure detection based on CNN considering the multi-channel characteristics of the original EEG signal. Gabeff et al. [24] focused on the performance of the model at the segment and seizure level based on the CNN method and discussed the correlation between these two levels for non-patient-specific seizure detection. Ein Shoka et al. first converted the EEG time series into a spectrograph image and fed to a CNN-based transform learning model, which showed superior performance than the common models [25]. KR et al. [26] used CNN to extract relevant features from the EEG signals. Shanmugam and Dharmar [27] proposed a hybrid 1D CNN-LSTM model that does not require feature extraction for the end-to-end seizure detection method.

Deep learning can also process time series data to better learn data correlations by preserving the state of the network. The CNN processes each element individually with no state saved between them. EEG signals are time series recordings. Therefore, learning the relationship between the entire sequence has a positive impact on the final classification of the signal. The recursive neural network (RNN) is a simple neural network with memory, which can maintain a state for all sequence elements. But it can only learn the dependencies of the sequences in the short term. As the number of layers increases, the network may have an untrainable problem, named the vanishing gradient.

The long short-term memory (LSTM) network is a variant of RNN, attaching a carrying information across multiple time steps when dealing with sequential problems. The LSTM can save information and prevent early information from fading away during later processing. In recent years, many studies have used LSTM to detect and predict seizures in EEG signals. Chakrabarti et al. [28] designed a simple LSTM for both invasive and noninvasive EEG recordings. The proposed method was effective in detecting epileptic seizures [28]. Hussein et al. [29] learn different patterns of EEG signals from a deep LSTM network, which can provide the high-level representations of EEG signals. Then, a dense layer was adopted to output the predicted results [29]. Hu et al. [30] proposed a novel method for seizure detection

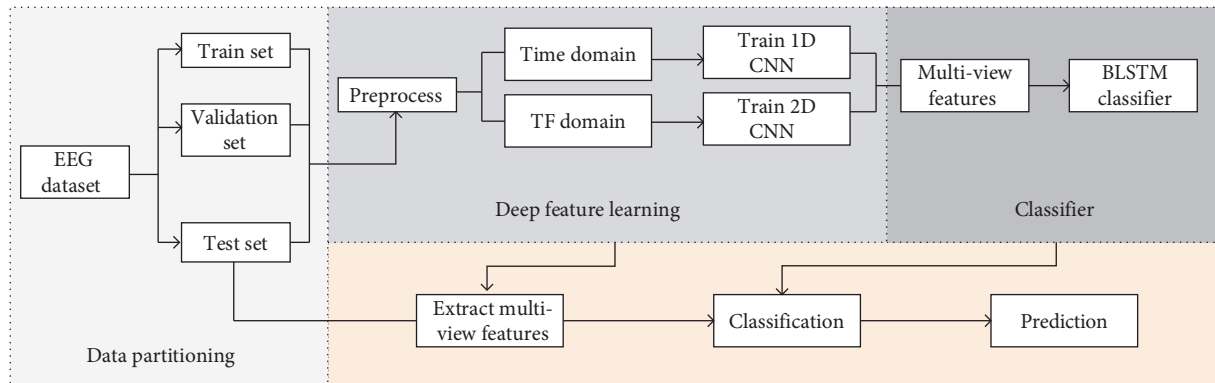


FIGURE 1: The flowchart of the proposed MDFLN for automatic seizure detection.

based on bidirectional long short-term memory (BLSTM) and introduced local mean decomposition to decrease the computational cost. A mixed model combining CNN and BLSTM was proposed to obtain the temporal evolution of seizure presentation from the EEG dataset with a small number of parameters [31]. The model they proposed only considered the time-domain features of EEG signals, which cannot provide efficient information. Hussain et al. [32] also proposed a hybrid model to decompose the raw EEG signal into time-domain, frequency-domain, and TF features to obtain enough detail information. But the CNN architecture they used is simple, it cannot learn high-level presentation of the EEG signals.

In this work, a hybrid network model is proposed for seizure detection (Figure 1). The model integrates the advantages of CNN and BLSTM. The designed CNN is used to deep learn the feature information of EEG signals. In order to obtain adequate feature information, multi-view features from time domain and TF domain are input into the 1D CNN and the 2D CNN to automatically extract features. 1D CNN is used to process the original EEG signals. 2D CNN is used to extract the TF image from the converted signals. The BLSTM network integrates the fine-grained feature vectors learned from the CNN structure to extract the long-term dependencies of epilepsy. Two-way learning enables the model to consider both past and future information of segmented signals, which will help detect seizures. The proposed MDFLN can imitate the clinical diagnosis process, where experts always mark the onset and end of seizures through the temporal evolution of signals.

Although previous work has also adopted CNN and BLSTM models for seizure detection, our model improves these methods on several important points. First, most previous studies focus on extracting single-view features. The previous work only considers time-domain features or 1D CNN to learn the temporal evolution. Accordingly, the proposed CNN operates on multi-view features that are extracted from the time series of the original signals and the TF image obtained by STFT transformation to obtain enough information. Second, when training the model, a validation set is used to prevent the model from overfitting. In addition, in order to improve the performance of the model, the BLSTM classification network is fine-tuned by combining the training set and

validation set. The effectiveness of the model is verified using two public datasets. The contributions of this work are presented as follows:

- (1) A MDFLN model combined CNN and BLSTM is proposed. The model uses CNN to perform deep feature learning on EEG signals and inputs the learned high-order features into the BLSTM for classification.
- (2) The multi-view feature information of the signal is comprehensively considered, including time-domain features and TF-domain features. The deep learning model based on multi-input features is constructed.
- (3) The BLSTM is used to classify the EEG signals, then the model is fine-tuned before the final classification. The performance of the model is verified in two datasets, which shows the effectiveness of the proposed method.

The content of this paper is organized as follows: in Section 2, the preprocessing method of the dataset is introduced. The proposed seizure detection method is developed. The experimental results and the analysis are presented in Section 3. Section 4 discusses comparisons with the state-of-the-art seizure detection methods using two public datasets. Finally, the conclusions are described in Section 5.

2. Methods and Materials

2.1. Dataset. The CHB-MIT dataset is a public dataset collected by Children's Hospital Boston, containing long-term, multi-channel EEG recordings from 23 neonatal epileptic patients with intractable epilepsy [33, 34]. Data acquisition was carried out by placing electrodes on the scalp of the patients in accordance with the International Standard 10–20 system. As shown in Table 1, this dataset collected records of 24 cases (chb01, chb02, ..., chb24) from 23 patients aged 1.5–22 years. The first 23 cases were from 22 patients with 17 women and 5 men. The sex and age of the 24th case was not informative. Each case contained a continuous file in the format of .edf from a single subject. Most of the files were digitized EEG signals that were an hour long recordings. In 24 cases, a total of 664 .edf files with 198 seizures were

TABLE 1: The information of the CHB-MIT database.

Patients	Gender age	Number of seizures (T_{\min} - T_{\max}) (s)	Total seizure time (s)	Total seizure-free time (s)
1	F-11	7 (28-102)	449	23,475
2	M-11	3 (10-83)	175	7,983
3	F-14	7 (48-70)	409	24,791
4	M-22	4 (49-116)	378	37,976
5	F-7	5 (97-121)	563	17,437
6	F-1.5	9 (13-21)	147	93,051
7	F-14.5	3 (87-144)	328	32,208
8	M-3.5	5 (135-265)	924	17,076
9	F-10	4 (63-80)	280	34,218
10	M-3	7 (36-90)	454	50,008
11	F-12	3 (23-753)	809	9,249
12	F-2	27 (14-98)	1,016	33,844
13	F-3	12 (18-71)	547	28,253
14	F-9	8 (15-42)	117	25,023
17	F-12	3 (89-116)	296	10,528
18	F-18	6 (31-69)	323	19,951
19	F-19	3 (78-82)	239	10,307
20	F-6	8 (30-50)	302	19,732
21	F-13	4 (13-82)	203	13,587
22	F-9	3 (59-75)	207	10,593
23	F-6	7 (21-114)	431	31,823
24	NR-NR	16 (17-71)	527	42,673

TABLE 2: The information of the Bonn database.

Subsets	Set A	Set B	Set C	Set D	Set E
Subjects	Healthy	Healthy	Epileptic patients	Epileptic patients	Epileptic patients
Patient Stage	Eye open	Eye close	Seizure free	Seizure free	Seizure activity
Electrode type	Surface	Surface	Intracranial	Intracranial	Intracranial

included. All EEG signals were sampled at 256 Hz with a 16-bit resolution. Most EEG signals contain 23 channels (24 or 26 in some cases). We could not read some of the channels in chb15 and chb16. Therefore, data from these two patients were removed.

The Bonn dataset was collected from the University of Bonn in Germany, which contains five subsets, represented as A, B, C, D, and E, respectively. Each subset contained 100 single-channel EEG segments with 23.6 s long. The sampling rate was 173.61 Hz with a resolution of 12 bits. All EEG signals were acquired with the 10-20 system [35]. As can be seen in Table 2, subsets A and B were from five healthy volunteers with eyes open and closed, respectively. Subsets C, D, and E were collected from five epileptic patients. Subsets C and D were the intervals during which the patients were in the stage of seizure-free. The data contained in E were in the stage of seizure activity.

2.2. Preprocessing of EEG Signals

2.2.1. Filtering. A fourth-order Butterworth filter was introduced. The 0.01–32 Hz frequency band was reserved for the diagnosis of epilepsy. The filter removed physiological

artifacts that confound epilepsy [36]. A non-overlapping window of 5 s was extracted from each channel. Each seizure activity of the recordings has been marked by clinical annotation to indicate the onset and the end of the seizure. Windows containing seizure events were labeled as positive samples. Consequently, windows that did not contain seizure events were labeled as negative samples.

2.2.2. Standardization. Feature scaling, also known as standardization, is a step in data preprocessing. This method converts the data value into a particular range. Feature scaling is a common data processing requirement for using Keras, Scikit learn, and deep learning. Generally speaking, it is unsafe to input relatively large data or heterogeneous data (for example, one feature of the data is in the range of 0–1 and the other is in the range of 100–200) into the neural network. It may result in a large gradient update. Therefore, the network cannot converge. To simplify the network, the value of the feature should be in the same range.

Standardization transforms the raw data into a distribution with a mean of 0 and a variance of Equation (1). The formula of standardization is as follows:

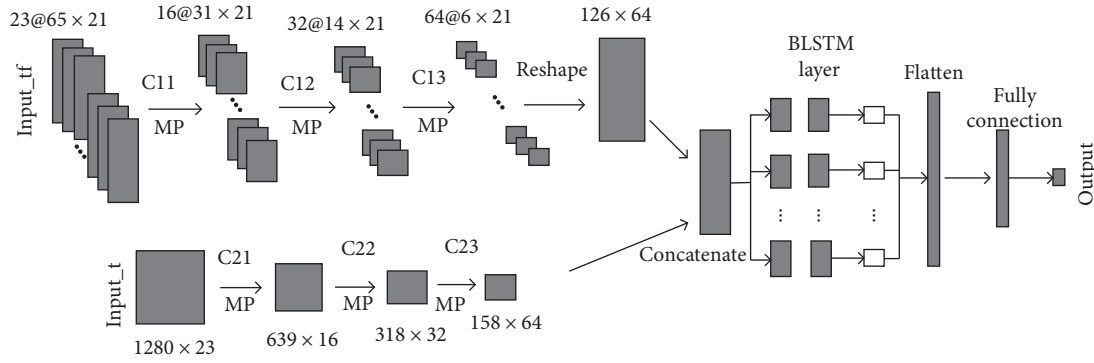


FIGURE 2: The detailed information of the proposed MDFLN model.

$$z = \frac{X - \mu}{\sigma}, \quad (1)$$

where z is the signal after standardization. X is the original value. μ and σ are the mean and variance of the recordings, respectively.

2.3. CNNs. CNNs perform well in computer vision problems using convolutional operations. It can make full use of the data and extract local features from the images for modular representation. Compared to the other image classification algorithms, CNN uses relatively less preprocessing. This means that the network optimizes the filters (or kernels) through automatic learning. In traditional algorithms, these filters are hand-designed. So, the major advantage of CNN is that they do not need prior knowledge or human intervention. This property not only makes neural networks outstanding for computer vision but also makes them particularly effective for signal processing.

A CNN consists of an input layer, a hidden layer, and an output layer. In a feed-forward neural network, intermediate layers are called hidden layers because their inputs and outputs are masked by activation functions and final convolutions. In CNNs, the hidden layers perform the dot product of the convolution kernel for the input matrix. As the convolution kernel slides along the input matrix, the convolution operation produces a feature map, which in turn contributes to the input of the next layer. This is followed by other layers, such as the pooling layer, the fully connected layer, and the normalization layer. According to the dimension of the input, two CNN functions are considered in this work, namely 1D CNN and 2D CNN, respectively. In 1D CNN, the kernel slides along one dimension, which can be used for processing such as time series data or natural language processing. In 2D CNN, the kernel slides along two dimensions through the data, which is usually used to deal with image data.

The CNNs uses three convolutional blocks to deeply learn the characteristics of the original signal during the process of extracting temporal features. As shown in Figure 2, each convolutional block includes the convolutional layer, the ReLU layer, the batch normalization layer, and the max pooling (MP) layer. This succession convolutional block extracts higher order features of the information of the

EEG signals. The convolutional layer is the main block in CNN, which performs a convolution operation on the input through the predefined filters and receptive field. The ReLU activation function $f(x) = \max(0, x)$ is applied to effectively remove negative values from the activation graph by setting them to zero. The ReLU layer introduces nonlinearity into the decision function and the whole network without affecting the receptive field of the convolutional layer. The batch normalization layer makes the input data conform to the same distribution with a simple and fast training process.

The pooling layer reduces the dimensionality of data by combining the output of one layer of clusters of neurons into individual neurons in the next layer. Pooling layer is used to gradually reduce the size of the representation space, reduce the number of parameters in the network, memory consumption and computation, so as to control overfitting. A very common form of maximum pooling is a layer with a size 2×2 filter, applied with a step length of Equation (2) as follows:

$$f_{X,Y} = \max_{a,b=0}^1 S_{2X+a,2Y+b}. \quad (2)$$

Each depth slice in the input along width and height is downsampled, discarding 75% of the data. Finally, the maximum value of the field is retained.

2.4. BLSTM. LSTM is a special RNN network, which can solve the problem of vanishing gradient in RNN and learn long-term dependence. But the one-way LSTM can only consider the forward relationship. The BLSTM takes into account the dependency between the front and back. Therefore, the output of BLSTM is more robust. The core idea of the LSTM model is the cell state, which is used to store the long-term state. Specifically, the LSTM uses gates to control the removal or addition of data in the cell state. Assume that x_t is the current input. h_{t-1} and c_{t-1} represent the short-term memory and the long-term memory at time $t-1$, respectively. Then the input gate I_t , the forgetting gate F_t , and the output gate O_t are defined as follows:

$$I_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (3)$$

$$F_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad (4)$$

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (5)$$

where W_i , W_f , W_o , b_i , b_f , and b_o are trainable model parameters. Then, a candidate memory cell is computed before updating the status,

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \quad (6)$$

where W_c and b_c are the trainable model parameters. According to the candidate memory cell \tilde{C}_t and long-term memory at time $t - 1$, the cell state C_t and h_t are updated as follows:

$$C_t = F_t \times C_{t-1} + I_t \times \tilde{C}_t, \quad (7)$$

$$h_t = O_t \times \tanh(C_t). \quad (8)$$

BLSTM is good at processing time series data. If the data are three-dimensional graphics such as images, it is difficult for BLSTM to describe this spatial feature because of the rich spatial information and the strong correlation between each point. Therefore, on the basis of BLSTM, convolution operation is added to capture spatial features, which will be more effective for image feature extraction.

After the CNN feature extraction phase, the feature vector of the segmented signal is classified as a series of binary predictive values. The BLSTM structure connects the outputs of two LSTMs. One that operates forward in the sequence, and the other that operates backward. Therefore, the BLSTM contains information from the past to the future at any given time. The bidirectional structure allows networks to learn the temporal evolution of the seizure. It ensures high-temporal resolution and low latency by using the entire recordings of the segmented signal in the network. Two BLSTM hide layers are used before the final prediction. The output is classified by the dense connection layer.

2.5. The Architecture of the Proposed MDFLN. Our model can be conceptualized as a multichannel feature extractor followed by a temporal detector (BLSTM). The outline of the network is shown in Figure 2. This illustration uses the CHB-MIT dataset to describe the network structure.

2.5.1. 1D CNN. For some time series problems, the effect of 1D CNN can be comparable to that of RNN. And the cost is usually much lower. 1D CNN extracts a patch from a sequence and applies a convolutional transformation to it, so patterns learned at one location in the sequence can also be recognized elsewhere. EEG signals in the time domain show how the values of the original signal change over time. Time-domain signals accurately reveal the location of the seizure. The signal analysis is performed based on the time–amplitude information of the signal components. However, these signals are not able to disclose the frequency range in which the spike occurs. The features of all 23 channels are integrated to form feature vectors. This feature

vector is used as the input of the model to train and classify the data.

In our work, 1D CNN uses the original EEG signal as input with segment length of 5 s. There are three convolution blocks, named C11, C12, and C13, respectively. Each block contains a convolutional layer with a ReLU layer, a batch normalization layer, and an MP layer. Specifically, the MP layer is not drawn, which is represented by MP in the figure. For C11, there are 16 filters with a size of 3. The result of convolution is input into a ReLU activation function followed by a MP layer over a region of 2. C12 and C13 use the same structure except that C12 and C13 have 32 and 64 filters, respectively.

2.5.2. 2D CNN. We consider converting the original signal to a TF map, an image-like format, to extract TF information. The time series EEG signals are usually transformed into image shapes using the wavelet transform and the Fourier transform [37, 38]. For seizure prediction and detection, they are an effective feature extraction method. In this paper, we use STFT to transform the raw EEG signal into a 2D matrix consisting of a frequency axis and a time axis. STFT is usually used to analyze nonstationary time-varying signals.

The TF-domain information of EEG signals is a coordinate system that evaluates the frequency characteristics of a signal by imitating the relationship between frequency and amplitude. The STFT method can convert the time-domain signal from each channel into the TF-domain signal. Figure 3 gives an example to transform the seizure and nonseizure signal into TF representation. The Hamming window is used and the length of the segment is 128. The TF domain representation of signals is suitable for nonstationary and time-varying signals, helping to extract temporal and spatial correlation information.

In the 2D CNN, the STFT transformation result of the 5 s window of the original signal is used as input. There are also three convolution blocks, named C21, C22, and C23, respectively. Each block contains a convolutional layer, a ReLU layer, a batch normalization layer, and an MP layer. For C21, there are 16 filters with a size of 3. The result of convolution is followed by an MP layer over a region of 2. C22 and C23 use the same structure except that C22 and C23 have 32 and 64 filters, respectively.

The window length is set to 5 s sampled at 256 Hz. So, the length of the segmented recordings is 1,280. The original input length of 1D CNN is $1,280 \times 23$, where 23 is the number of channels. The Hamming window is used for transformations with a length of 128, so the input of 2D CNN is $65 \times 21 \times 23$. Table 3 presents the layer information of the designed network and shows the output size of each layer. Furthermore, to mitigate the imbalance in the dataset, we overlapped the samples in the ictal period and did not overlap the samples in the interictal period.

As in VGGNet, we double the number of channels per block after the pooling layer. This process prevents information from being lost and ensures that each convolutional block requires approximately the same amount of computation. Global average pooling is used to represent the last layer

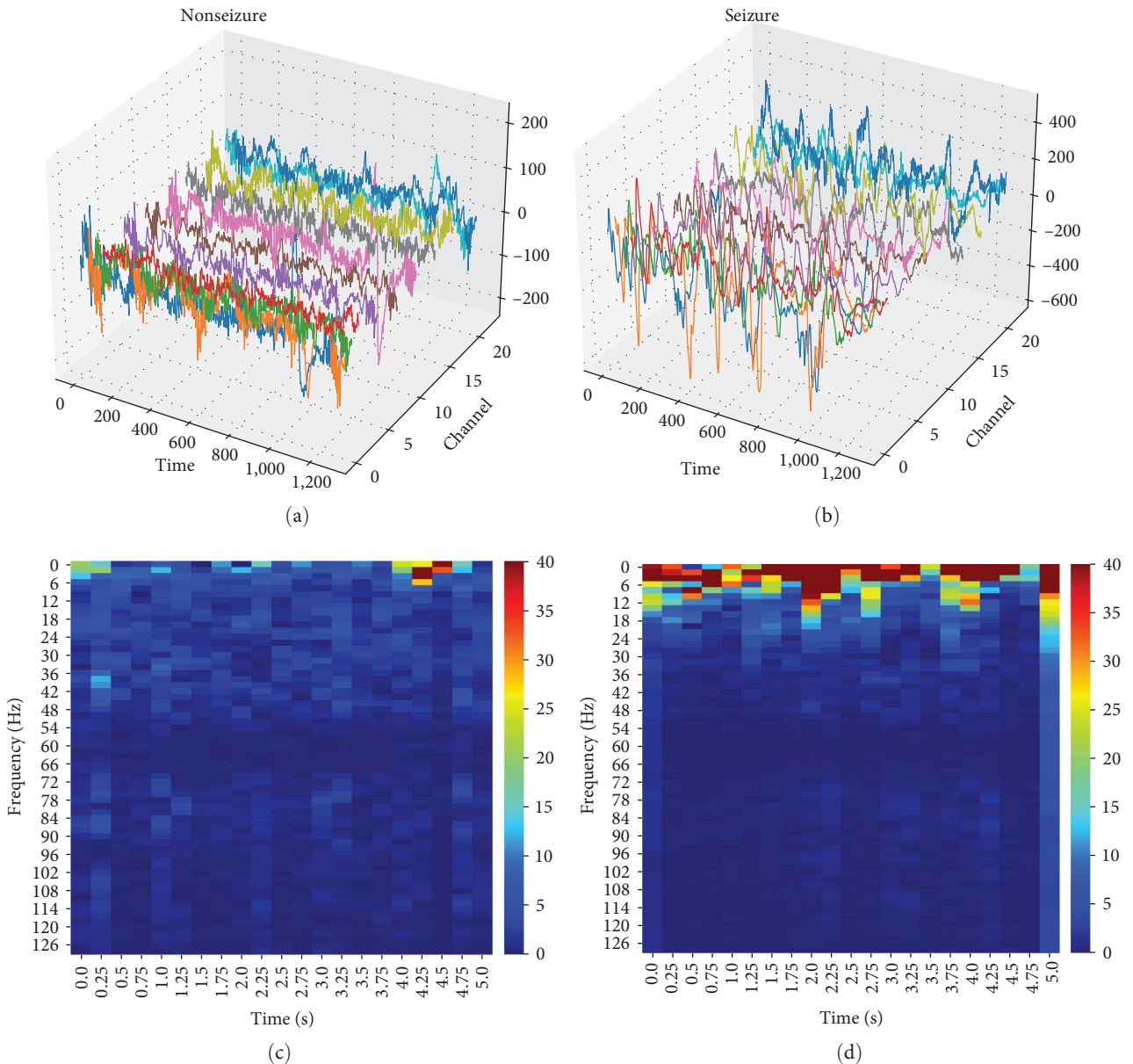


FIGURE 3: Examples of STFT of 5-s window. The raw nonseizure (a) and seizure (b) EEG signals. The TF representation of nonseizure (c) and seizure (d) signals.

of convolutional blocks. The output of each kernel is averaged so that each kernel has a single feature. This procedure has a regularizing effect on the network; broadly, it reduces overfitting, as the subsequent recurrent layers receive information pooled across the entire one second window, thus mitigating overfitting to isolated data irregularities.

CNN and BLSTM are effective tools for dealing with time and noise signals and can achieve efficient and stable classification accuracy. CNN is a deep learning algorithm that can effectively classify data into multiple categories by automatically learning features. It is generally insensitive to noise and can gather valuable information in the presence of noise which can be regarded as a feature extractor. BLSTM can be a classifier to identify seizure events in EEG signals, but it

contains too much redundant information, resulting in high-time consumption. Therefore, a hybrid CNN-BLSTM model is proposed to classify EEG signals. The model can effectively use the time dependence in time series to detect seizures. Instead of getting all types of features manually, the time-domain or TF-domain signal are input directly. At the beginning of the model, CNN is used to obtain reliable and distinguishable features, consisting of a convolution layer, a linear unit Relu, an MP layer, and one or more fully connected layers. The feature learned by deep learning becomes more abstract. The convolution layer contains K filters of size 3. The result of the ReLU activation function is regarded as the output of the CNN layer. A nonlinear operation is performed in the network to replace negative output with 0. The output of this layer has the same

TABLE 3: Layers of the designed network.

	Time domain		TF domain	
	Layer info	Output size	Layer info	Output size
Input	—	(1,280, 23)	—	(65, 21, 23)
C11/C21	Conv1D(3) × 16	(1,278, 16)	Conv2D(3, 1) × 16	(63, 21, 16)
	ReLU	(1,278, 16)	ReLU	(63, 21, 16)
	BatchNor	(1,278, 16)	BatchNor	(63, 21, 16)
	MaxP, s:2	(639, 16)	MaxP(2, 1), s:2	(31, 21, 16)
C12/C22	Conv1D(3) × 32	(637, 32)	Conv2D(3, 1) × 32	(29, 21, 32)
	ReLU	(637, 32)	ReLU	(29, 21, 32)
	BatchNor	(637, 32)	BatchNor	(29, 21, 32)
	MaxP, s:2	(318, 32)	MaxP(2, 1), s:2	(14, 21, 32)
C13/C23	Conv1D(3) × 64	(316, 64)	Conv2D(3, 1) × 64	(12, 21, 64)
	ReLU	(316, 64)	ReLU	(12, 21, 64)
	BatchNor	(316, 64)	BatchNor	(12, 21, 64)
	MaxP, s:2	(158, 64)	MaxP(2, 1), s:2	(6, 21, 64)
	—	—	Reshape	(126, 64)
Concatenate		(284, 64)		
BLSTM 1		(284, 40)		
BLSTM 2		(284, 40)		
Flatten		11,360		
Dropout		11,360		
Dense		128		
Dense		1		

size as the input. The MP layer converts the input data that is the same size as the kernel into a single output with the maximum number observed.

RMSProp algorithm is used to determine the optimal weights and bias set of the neural network, which significantly reduces the loss function. RMSProp uses an exponentially weighted moving average to speed up the optimization process. The binary cross-entropy loss function is used to train the model. The learning rate for the function to move through the search space is set to 0.0001. A lower learning rate leads to a more consistent result, which also leads to more training time. The learned features are transferred by an activation function, SoftMax or Sigmoid, to obtain the probability of the input. After the input passes through the network, the result is reduced and downsampled. The proposed network is designed to alleviate overfitting and can handle LSTM models, performing quite well in time series classification tasks. The classifier consists of two BLSTM layers and a fully connected layer, which outputs the results of the seizure detection. The result of the classifier is a probability value of the input signal that matches a specific type of seizure. In particular, the proposed architecture trains and fine-tunes the top-level classification portion of the network on the two datasets. Compared to the advanced methods based on machine learning and deep learning, the proposed method demonstrates effectiveness and robustness by providing the most impressive performance in seizure recognition. The proposed technique is also believed to maintain stable efficiency in the presence of certain EEG artifacts

and environmental disturbances, which is more suitable for the clinical diagnosis.

2.6. Metrics. The measurement metrics used in this work are accuracy, specificity, sensitivity, and area under curve (AUC). Accuracy, defined as the proportion of samples that are correctly classified, is the most common performance metric. Specificity, also known as the true negative rate, refers to the proportion of nonseizure samples that are correctly classified. Sensitivity, also known as recall, is the proportion of positive samples in the original sample that are correctly predicted. The classifier generates a probability prediction for each sample during classification and compares this prediction value with a threshold. If the prediction value is greater than the threshold, it is positive; otherwise, it is negative. The ROC curve does not specify a fixed threshold, but tries all possible thresholds (cut-off points) and calculates a pair of true positive rate (TPR) and false positive rate (FPR) at each possible threshold. The AUC was used as an indicator of the model. The higher the AUC value, the higher the accuracy. Performance metrics are defined as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}} \times 100\%, \quad (9)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100\%, \quad (10)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\%, \quad (11)$$

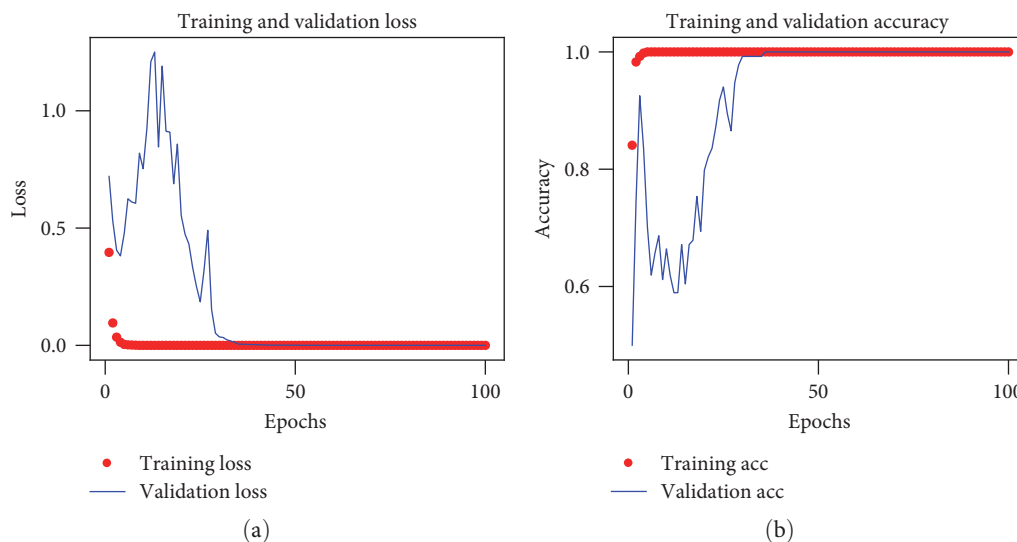


FIGURE 4: (a) The loss and (b) accuracy of the MDFLN during training and validation.

where TP represents the number of positive samples correctly predicted. TN represents the number of negative samples that are correctly predicted. FN represents the number of positive samples predicted to be negative samples. FP represents the number of negative samples predicted to be positive.

3. Results

3.1. Experimental Configuration. The proposed seizure detection network model is implemented by using PYTHON 3.9 on a Thinkpad T14, Intel i5 10th, and RAM 16 G. The Python program uses the Keras framework and is implemented with TensorFlow as the backend. The specific parameter settings and the execution process in the code have been loaded on GitHub: <https://github.com/chloeqisun/MDFLN>.

The proposed model is validated and tested on raw EEG signals from two public datasets. The performance of the model is evaluated with metrics including accuracy, specificity, sensitivity, and AUC. The different length of the segments has also tried to select suitable length of segments. Furthermore, fine tuning was used to tune the model.

The EEG signals are classified as seizure events or non-seizure events on the CHB-MIT dataset. Three different classification cases are distinguished on the Bonn dataset. Case 1: Sets A, B, C, and D combine as normal class; Set E is the epilepsy class. Case 2: Sets A and B are healthy persons; Sets C and D are the interictal period of patients with epilepsy; Set E is the ictal period of seizures. Case 3: Sets A, B, C, D, and E are classified into one category, respectively.

The EEG signal of a patient is a long time series that needs to be segmented before analyzing. In order to get a suitable segment length, we try different segment lengths in the range of [1,10] with a step size of 1. Through training and testing on the dataset of the first patient in CHB-MIT, it is concluded that the segment with a length of 5 s has better classification performance. The classification result improves with increasing segment length at first. When the segment

length is greater than 5 s, the classification performance does not improve with the increase of segment length. Therefore, during data preprocessing, the segment length is set to 5 s. The feature information contained in different lengths of segments is varied. Short segmented signal recordings may not capture the time evolution information of the signal. However, longer segments are not more helpful in detecting the collected seizure signals. For the Bonn dataset, the length of the segment is also set to 5 s.

3.2. Seizure Detection Results

3.2.1. Base Network. In this work, the fivefold cross-validation method is used to validate the proposed model. First, we divide the data into training set and testing set. In order to avoid over-fitting, a part of the training set is divided into the validation set. The remaining part of the training set is used to train the model. And the trained model is saved. Deep learning methods can perform gradient optimization, and thus can help the algorithm converge to a global optimum. The validation accuracy and validation loss curve can reflect the classification performance of the model. In this work, the convergence of the loss function is studied to analyze the performance of the proposed model. Figure 4(a) shows the convergence curve of the loss function. It can be seen that the objective function remains the same after several epochs. After 30 rounds, the loss function converges and no improvement in terms of accuracy, depicted in Figure 4(b), occurs. Therefore, in our experimental setting, the EarlyStopping callback function of Keras is used to control the learning of the model. When the classification accuracy of the training does not improve in 10 rounds, the training is stopped in advance. And the model with the least loss function is recorded.

The classification results for the CHB-MIT dataset are given in Table 4 which contains EEG data for 22 patients. For each patient, a model needs to be created. It can be seen from the table that the models for most patients have good classification results. The multi-input deep learning model can

TABLE 4: The classification results of the proposed MDFLN based on two public datasets.

Dataset	PID	Base network				Fine-tuning network				
		Accuracy	Specificity	Sensitivity	AUC	Accuracy	Specificity	Sensitivity	AUC	
CHB-MIT	1	99.76	99.52	100	100	100	100	100	100	
	2	99.06	98.75	99.38	98.63	100	100	100	100	
	3	95.71	93.82	97.6	97.01	98.52	97.31	99.73	97.74	
	4	94.34	91.72	96.97	99.13	97.11	96.14	98.08	99.11	
	5	94.88	91.42	98.33	98.79	97.49	96.66	98.33	99.82	
	6	91.56	98.26	84.86	94.42	97.75	97.27	98.22	99.17	
	7	98.23	99.37	97.1	98.54	97.91	99.05	96.78	98.67	
	8	97.78	97	98.55	99.65	97.66	97	98.33	99.61	
	9	98.65	97.69	99.62	99.22	98.46	97.31	99.62	99.37	
	10	99.15	99.04	99.28	99.23	96.9	96.9	96.9	99.03	
	11	96.14	93.8	98.48	90.69	98.04	97.98	98.1	95.8	
	12	96.6	96.48	96.72	99.49	98.53	98.98	98.07	99.75	
	13	98.38	98.25	98.5	99.88	99.13	99.5	98.75	99.85	
	14	96.37	97.8	94.95	97.12	99.27	98.54	100	100	
	17	96.8	98.58	95.03	98.92	96.62	99.64	93.59	97.03	
	18	97.79	95.91	99.66	99.76	99.15	99.66	98.64	99.79	
	19	96.4	94.59	98.21	94.89	96.18	94.15	98.22	95.51	
	20	98.29	97.72	98.87	99.74	96.4	93.96	98.85	99.82	
	21	97	96.2	97.81	97.99	96.43	93.96	98.9	94.6	
	22	96.66	94.86	98.46	100	99.74	99.49	100	100	
	23	98	96.74	99.25	99.92	98.99	99.24	98.74	99.7	
	24	98.22	98.44	98	99.73	97.77	98.21	97.32	99.27	
		Average	97.08	96.63	97.53	98.31	98.09	97.77	98.42	98.8
	Bonn	Case 1	98.32	99.03	95.36	99.14	98.4	98.71	97.02	99.07
Case 2		92.91	94.46	89.36	—	93.12	94.84	89.68	—	
Case 3		89.76	93.6	74.4	—	91.28	94.55	78.2	—	

effectively classify EEG signals. The average classification accuracy of these patients is 97.08%. For the Bonn dataset, the accuracy of the three cases is 98.32%, 92.91%, and 89.76%, respectively. In the process of evaluating the performance of the model, the patient dataset is divided into a training set, a validation set, and a testing set. The model is based on the training set. The validation set is used to supervise the training process to prevent the model from overfitting. This method reduces the data in the training set, thus affecting the performance of the model. To build the model with more training data, the model is fine-tuned by combining the training set and the validation set of the patients.

3.2.2. Fine-Tuned Network. To improve the performance of the model, we fine-tune the model (Figure 5). The CNN feature extraction part of the initial model is frozen. The BLSTM classification layer is fine-tuned. Before testing, the training set and the validation set are combined. Then, the model is retrained by fine-tuning the saved model. The fine-tuned model is used to calculate the classification performance on the testing set. Specifically, the CNN layer is frozen. The BLSTM and dense layers are trained to obtain a fine-tuned model. Table 4 shows the classification based on two datasets. As can be seen from the table, the average values of accuracy, sensitivity, and specificity

are improved after the model is fine-tuned. The classification performance shows improvement in 13 of the 22 patient models. This happens because when fine-tuning the network, the training set and the verification set are combined to train the model, which increases the training set to deep learn the features. The frozen CNN layer is equivalent to the automatic feature learning process of the original signals. And the parameters of the BLSTM are adjusted in the merged training set. The results of all patients of the fine-tuned model are above 96% with accuracy. The AUC of the model classification is also higher, which shows the validity of the model. The classification results of the Bonn dataset are presented in Table 4. As you can see from the table, the performance of the model is significantly improved after fine-tuning. The performance of the binary classification is also improved effectively.

3.2.3. Multi-Input Features Learning Analysis. To verify the effect of single-view and multi-view features on the performance of the model, we compare the difference between them in Figure 6. The time-domain learning model just used 1D CNN to learn the features of EEG signals. The TF-domain learning model used 2D CNN to obtain features. These two single-input models both used BLSTM network to classify the signals.

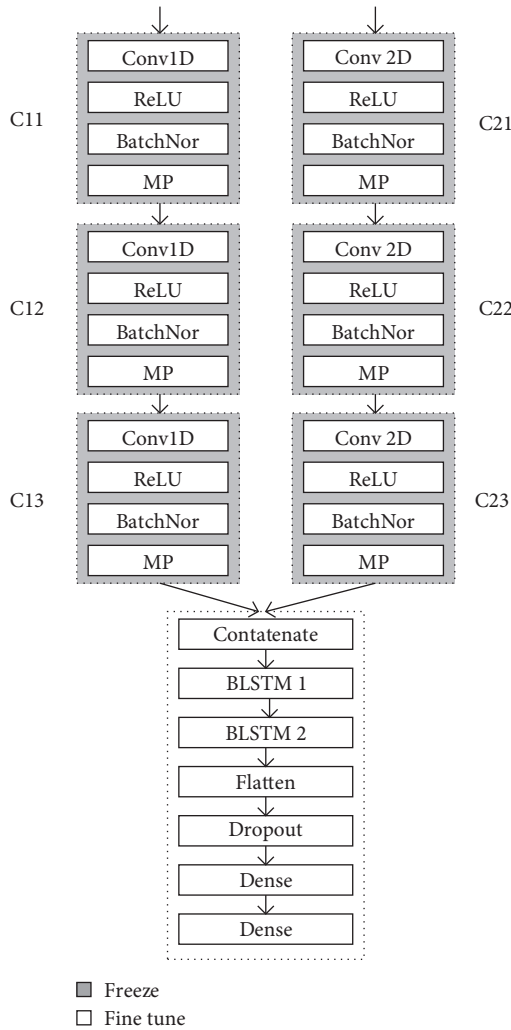


FIGURE 5: Fine-tuning the last module of the network.

On the CHB-MIT dataset (Figure 6(a)), the classification results of the time domain and the TF domain are 89.67% and 80.04% with average accuracy, respectively. Multi-input networks achieve better results for each patient. The 1D CNN, which takes into account the temporal information of EEG signals, performs better in identifying seizure events compared to extracting spatial information from the TF image. We observed that in the 2D CNN with BLSTM, the EEG signals of some patients cannot be effectively identified, leading to classification results lower than baseline. And on the Bonn dataset (Figure 6(b)), the classification results of the time domain and the TF domain are 92.24% and 74.87% with average accuracy for three cases, respectively. This result is consistent with the findings from the CHB-MIT dataset. When considering only the use of TF-domain information, the classification performance is the poorest. However, by combining the information learned by 1D CNN and 2D CNN, the classification performance is effectively improved. Specifically, the classification accuracy of the proposed method based on multi-view features is at least 2.2% higher than single-view features (CHB-MIT, 9.39%; Bonn, 2.2%).

When only a single input is considered, its performance is far less good than that of multiple inputs. The features of the time domain are more beneficial for the classification of EEG signals than those of the TF domain. The time-domain feature considers deep learning of segmented signals and uses BLSTM networks to learn the temporal development patterns of EEG signals. The experimental results show its effectiveness in the recognition of EEG signals. The frequency features of the signals also reflect different patterns of activity in the brain. Detecting seizure in EEG signals only with its temporal characteristics will degrade classification performance. Therefore, when training the networks, we also integrate the TF information of the signal into the training set. When EEG classification is performed using time-domain and TF-domain features of signals, the classification results for accuracy, specificity, and sensitivity are high. However, the classification performance of a single feature information is far inferior to the classification effect of the multi-input model. Comprehensively, considering the multi-view features of signals can help to improve the accuracy of recognition for the EEG signals.

4. Discussions

To evaluate the performance of the proposed method, a comparison with the state-of-the-art literature using the same dataset is presented in Table 5. For the CHB-MIT dataset, the comparison results show that the proposed MDFLN model can recognize EEG signals successfully. The multi-input deep feature learning method achieves the highest classification performance with a sensitivity of 98.42% in all mentioned works. Zhao et al. [46] proposed a CNN + Transformer method with the highest accuracy. And the transfer learning method was used by Ein Shoka et al. [25] for sensitivity, specificity, and accuracy are 88.89%, 84.21%, and 86.11%, respectively. For the Bonn dataset, the specificity obtained by the proposed method is superior to all works [39, 40, 43, 47, 49, 50]. Li et al. [39] designed a unified temporal-spectral squeeze-and-excitation network for classification task, achieving an accuracy of 99.8%. Different feature scaling techniques were employed by Thara et al. [47] to find the best results with a sensitivity of 98.59%. When comparing with existing literature, the comparison on the Bonn dataset only involves the comparison of results for the binary classification. Specifically, sets A, B, C, and D represent samples with no seizure events, and E represents samples with seizure events. The results of the proposed method are also acceptable, since there is a 97.02% value of sensitivity and a quite high accuracy of 98.4% for case 1. Although the proposed method is behind part of the studies summarized in Table 5 considering the accuracy metric, it is also efficient in the identification of epileptic seizures in EEG signals.

The MDFLN model is channel independent, as it takes into account entire channels without channel selection. This section uses the same data for the proposed method as mentioned in other literature. The related methods of the dataset are compared. The methods that simultaneously consider the time and frequency information of the signal are fewer. The proposed MDFLN simultaneously considers the time and

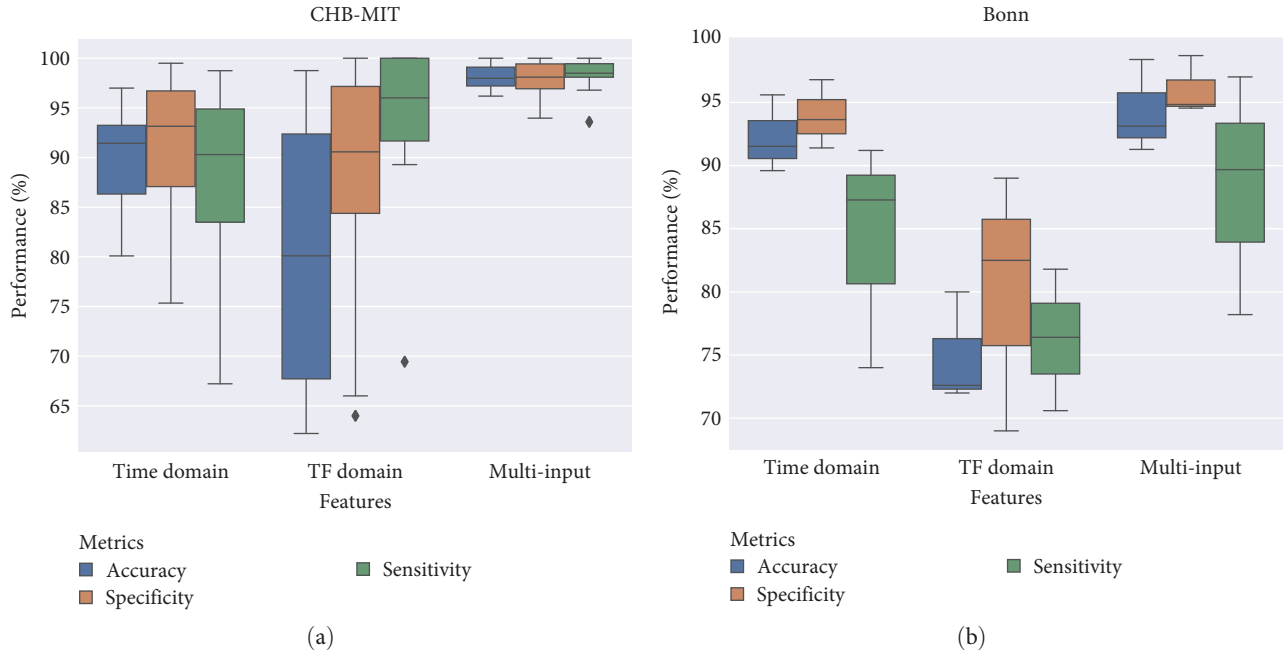


FIGURE 6: The comparison of classification performance between single-view and multi-view features: (a) CHB-MIT dataset and (b) Bonn dataset.

TABLE 5: Comparisons with state-of-the-art seizure detection methods using the same dataset.

Authors	Methodology	Dataset	Sensitivity(%)–specificity(%)–accuracy(%)
Truong et al. [21]	STFT, CNN	CHB-MIT	81.2–NR–NR
Tian et al. [22]	Deep multi-view feature learning, FFT, WPD, CNN	CHB-MIT	96.7–99.1–98.3
Li et al. [39]	Unified temporal-spectral squeeze-and-excitation network	CHB-MIT	92.41–96.05–95.96
	—	Bonn	NR–NR–99.8
Peng et al. [40]	Stein kernel-based sparse representation	CHB-MIT	97.85–98.57–98.21
	—	Bonn	98.43–98.67–98.67
Deng et al. [41]	Transductive transfer learning fuzzy system	CHB-MIT	97.16–97.03–97.15
Hossain et al. [42]	Raw EEG as 2D Array, CNN	CHB-MIT	90–91.65–98.05
Shoeibi et al. [43]	Handcrafted features, SVM, KNN, CNN	Bonn	NR–NR–99.53
Li and Chen [44]	FFT, 2D matrix, SVM	CHB-MIT	98.28–98.5–98.47
Jiang et al. [45]	Personal correlation coefficient and mutual information, SVM	CHB-MIT	97.72–95.62–96.67
Zhao et al. [46]	CNN, transformer	CHB-MIT	97.70–97.6–98.76
Thara et al. [47]	DNN, Different feature scaling techniques	Bonn	98.59–91.47–97.21
Xiong et al. [48]	Multivariate variational mode decomposition, RF	CHB-MIT	98.24–97.83–97.39
Ein Shoka et al. [25]	Transfer learning, CNN	CHB-MIT	88.89–84.21–86.11
Akyol [49]	Stacking ensemble approach, deep neural networks (DNN)	Bonn	93.11–98.18–97.17
Pan et al. [50]	Hybrid input, DWT, STFT, FFT, CNN	Bonn	NR–NR–97.89
This work	Multi-view features, MDLFL	CHB-MIT	98.42–97.77–98.09
	—	Bonn	97.02–98.71–98.4

The NR stands for not reported values.

frequency features of the signal to construct a robust and comprehensive system to efficiently detect seizure events. The previous CNN network is the process of feature extraction. The subsequent BLSTM network is a classification network. The CNN feature extraction part extracts time and frequency features that are input into BLSTM as a feature vector for seizure detection.

In general, research on EEG signal classification tends to use a single dataset to design a framework. However, this work uses two datasets to verify the classification performance of the model. Not only for binary classification problems, multi-classification problems can also train effective models. Deep learning networks can essentially learn features from a dataset. Taking advantage of this deep learning property, the proposed method can effectively avoid the time-consuming feature extraction process. Once the model is trained, the model can be saved, and unseen EEG data can be detected and analyzed. This could help to improve the quality of life of patients with intractable epilepsy.

The development of deep learning networks is an important advance in machine learning. CNNs can potentially learn features without additional feature extraction processes. The proposed CNN network includes convolutional layers, MP layers, and batch normalization. The original signals from multiple channels are used as the CNN input with linear filters. The MP layer downsamples the data with the pooling size 2, which reduces the dimensionality of the data with minimal loss. To avoid overfitting, dropout and batch normalization are used to fine-tune the network.

The learning ability of LSTM is used to evaluate the classification performance of features in decision-making. Traditional RNN has the problem of vanishing gradient, which makes it impossible to realize long sequence memory. Compared with RNN, LSTM changes the structure of the internal computing network and adds a memory unit. The memory unit is used to store the useful content of the previous sequence and apply it to the later sequence, which solves the problem that the RNN network cannot realize long sequence memory. The nature of bidirectional networks is that they use past and future information to perform classification in EEG segments. CNN learns spatial information in the receptive field. With the deepening of network learning, the features of CNN learning become more and more abstract. For EEG seizure detection, CNN must learn abstract spatial features. For time series EEG signals, the longer temporal dependencies obtained by deep learning are helpful for seizure detection. Therefore, the CNN and BLSTM structures are used to make the detection results more robust. Although the proposed MDLFLN effectively realizes seizure detection, there are some limitations in this work:

- (1) For the CHB-MIT dataset, short-term EEG signals are selected to train the model. But in long-term EEG signals, the dataset is imbalanced between ictal and interictal recordings. So, the proposed method should handle the problem of the imbalance of dataset.
- (2) A more generalized model needs to be designed. The model we used to detect seizures is patient-specific,

which does not generalize to the different patient patterns. Therefore, transfer learning is an effective method to establish a cross-patient model.

- (3) The proposed method utilizes information from all channels, which does not capture the mutual influence relationships between channels. Signal modeling approaches based on graph theory can capture directed or undirected influence relationships, providing a more accurate identification of the seizure events.

5. Conclusions

Epilepsy is one of the most common and extremely harmful neurological diseases. Some epileptic seizures can be treated with medication. However, there are some refractory seizures that cannot be precisely searched for the epileptogenic zone and accurately classified or diagnosed. To formulate personalized seizure prediction and treatment strategies, an automatic seizure detection method based on MDLFLN is proposed. We investigate multi-input deep learning models to extract feature information from the original and transformed signals. The results show that the classification performance of the multi-input model is much higher than that of the single-input model. The proposed model comprehensively considers the time-domain and TF-domain features of the signal and effectively achieves the identification of seizure events and nonseizure events. Furthermore, the classification performance is improved by fine-tuning the model. Finally, the effectiveness of the automatic seizure classification method is verified in two public datasets, which can help the experts with clinical treatment.

Data Availability

The public CHB-MIT database provided by the Boston Children's Hospital is used in this paper. It can be found in the website: <https://physionet.org/physiobank/database/chbmit/>. The Bonn database can be found in the website: <http://www.meb.uni-bonn.de/epileptologie/science/physik/eegdata.html>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the Postgraduate Research & Practice Innovation Program of Jiangsu Province in 2021 under grant no. KYCX21_0718, and the Natural Science Research Start-up Foundation of Recruiting Talents of Nanjing University of Posts and Telecommunications under grant no. NY222059.

References

- [1] A. T. Tzallas, M. G. Tsipouras, D. G. Tsalikakis et al., "Automated epileptic seizure detection methods: a review study," in *Epilepsy-Histological, Electroencephalographic and Psychological Aspects*, pp. 75–98, IntechOpen, 2012.

- [2] A. Shoeibi, N. Ghassemi, M. Khodatars et al., "Detection of epileptic seizures on EEG signals using ANFIS classifier, autoencoders and fuzzy entropies," *Biomedical Signal Processing and Control*, vol. 73, Article ID 103417, 2022.
- [3] U. R. Acharya, Y. Hagiwara, and H. Adeli, "Automated seizure prediction," *Epilepsy & Behavior*, vol. 88, pp. 251–261, 2018.
- [4] 2021, https://commons.wikimedia.org/wiki/File:Epilepsy_world_map_-_DALY_-_WHO2002.svg.
- [5] S. M. Usman, S. Khalid, R. Akhtar, Z. Bortolotto, Z. Bashir, and H. Qiu, "Using scalp EEG and intracranial EEG signals for predicting epileptic seizures: review of available methodologies," *Seizure*, vol. 71, pp. 258–269, 2019.
- [6] S. Selim, E. Elhinamy, H. Othman, W. Abouelsaadat, and M. A.-M. Salem, "A review of machine learning approaches for epileptic seizure prediction," in *2019 14th International Conference on Computer Engineering and Systems (ICCES)*, pp. 239–244, IEEE, Cairo, Egypt, December 2019.
- [7] M. Natu, M. Bachute, S. Gite, K. Kotecha, and A. Vidyarthi, "Review on epileptic seizure prediction: machine learning and deep learning approaches," *Computational and Mathematical Methods in Medicine*, vol. 2022, Article ID 7751263, 17 pages, 2022.
- [8] A. Emami, N. Kunii, T. Matsuo, T. Shinozaki, K. Kawai, and H. Takahashi, "Seizure detection by convolutional neural network-based analysis of scalp electroencephalography plot images," *NeuroImage: Clinica*, vol. 22, Article ID 101684, 2019.
- [9] A. Temko, E. Thomas, W. Marnane, G. Lightbody, and G. Boylan, "EEG-based neonatal seizure detection with support vector machines," *Clinical Neurophysiology*, vol. 122, no. 3, pp. 464–473, 2011.
- [10] L. Logesparan, A. J. Casson, and E. Rodriguez-Villegas, "Optimal features for online seizure detection," *Medical & Biological Engineering & Computing*, vol. 50, pp. 659–669, 2012.
- [11] E. Pippa, E. I. Zacharakis, I. Mporas et al., "Improving classification of epileptic and non-epileptic EEG events by feature selection," *Neurocomputing*, vol. 171, pp. 576–585, 2016.
- [12] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Automatic seizure detection based on time–frequency analysis and artificial neural networks," *Computational Intelligence and Neuroscience*, vol. 2007, Article ID 80510, 13 pages, 2007.
- [13] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Epileptic seizure detection in EEGs using time–frequency analysis," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 703–710, 2009.
- [14] 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.
- [15] T. T. Chowdhury, S. A. Fattah, and C. Shahnaz, "Seizure activity classification based on bimodal Gaussian modeling of the gamma and theta band IMFs of EEG signals," *Biomedical Signal Processing and Control*, vol. 64, Article ID 102273, 2021.
- [16] M. Zabihi, S. Kiranyaz, A. B. Rad, A. K. Katsaggelos, M. Gabbouj, and T. Ince, "Analysis of high-dimensional phase space via Poincaré section for patient-specific seizure detection," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 3, pp. 386–398, 2016.
- [17] Y. Jiang, W. Chen, and M. Li, "Symplectic geometry decomposition-based features for automatic epileptic seizure detection," *Computers in Biology and Medicine*, vol. 116, Article ID 103549, 2020.
- [18] A. Subasi, J. Kevric, and M. Abdullah Canbaz, "Epileptic seizure detection using hybrid machine learning methods," *Neural Computing and Applications*, vol. 31, pp. 317–325, 2019.
- [19] Q. Sun, Y. Liu, S. Li, and C. Wang, "Automatic epileptic seizure detection using PSO-based feature selection and multilevel spectral analysis for EEG signals," *Journal of Sensors*, vol. 2022, Article ID 6585800, 16 pages, 2022.
- [20] A. Shoeibi, P. Moridian, M. Khodatars et al., "An overview of deep learning techniques for epileptic seizures detection and prediction based on neuroimaging modalities: methods, challenges, and future works," *Computers in Biology and Medicine*, vol. 149, Article ID 106053, 2022.
- [21] N. D. Truong, A. D. Nguyen, L. Kuhlmann et al., "Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram," *Neural Networks*, vol. 105, pp. 104–111, 2018.
- [22] X. Tian, Z. Deng, W. Ying et al., "Deep multi-view feature learning for EEG-based epileptic seizure detection," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 10, pp. 1962–1972, 2019.
- [23] A. O'Shea, G. Lightbody, G. Boylan, and A. Temko, "Neonatal seizure detection from raw multi-channel EEG using a fully convolutional architecture," *Neural Networks*, vol. 123, pp. 12–25, 2020.
- [24] V. Gabeff, T. Teijeiro, M. Zapater et al., "Interpreting deep learning models for epileptic seizure detection on EEG signals," *Artificial Intelligence in Medicine*, vol. 117, Article ID 102084, 2021.
- [25] A. A. Ein Shoka, M. M. Dessouky, A. El-Sayed, and E. El-Din Hemdan, "An efficient CNN based epileptic seizures detection framework using encrypted EEG signals for secure telemedicine applications," *Alexandria Engineering Journal*, vol. 65, pp. 399–412, 2023.
- [26] A. B. KR, S. Srinivasan, S. K. Mathivanan et al., "A multi-dimensional hybrid CNN-BiLSTM framework for epileptic seizure detection using electroencephalogram signal scrutiny," *Systems and Soft Computing*, vol. 5, Article ID 200062, 2023.
- [27] S. Shanmugam and S. Dharmar, "A CNN-LSTM hybrid network for automatic seizure detection in EEG signals," *Neural Computing and Applications*, vol. 35, pp. 20605–20617, 2023.
- [28] S. Chakrabarti, A. Swetapadma, and P. K. Pattnaik, "A channel independent generalized seizure detection method for pediatric epileptic seizures," *Computer Methods and Programs in Biomedicine*, vol. 209, Article ID 106335, 2021.
- [29] R. Hussein, H. Palangi, R. K. Ward, and Z. J. Wang, "Optimized deep neural network architecture for robust detection of epileptic seizures using EEG signals," *Clinical Neurophysiology*, vol. 130, no. 1, pp. 25–37, 2019.
- [30] X. Hu, S. Yuan, F. Xu, Y. Leng, K. Yuan, and Q. Yuan, "Scalp EEG classification using deep Bi-LSTM network for seizure detection," *Computers in Biology and Medicine*, vol. 124, Article ID 103919, 2020.
- [31] J. Craley, E. Johnson, C. Jouny, and A. Venkataraman, "Automated inter-patient seizure detection using multichannel convolutional and recurrent neural networks," *Biomedical Signal Processing and Control*, vol. 64, Article ID 102360, 2021.
- [32] W. Hussain, M. T. Sadiq, S. Siuly, and A. U. Rehman, "Epileptic seizure detection using 1 D-convolutional long

- short-term memory neural networks,” *Applied Acoustics*, vol. 177, Article ID 107941, 2021.
- [33] A. L. Goldberger, L. A. N. 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.
- [34] A. H. Shoeb, *Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment*, Massachusetts Institute of Technology, 2009.
- [35] 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 ID 061907, 2001.
- [36] L. S. Vidyaratne and K. M. Iftekharuddin, “Real-time epileptic seizure detection using EEG,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 11, pp. 2146–2156, 2017.
- [37] B. H. Brinkmann, J. Wagenaar, D. Abbot et al., “Crowdsourcing reproducible seizure forecasting in human and canine epilepsy,” *Brain*, vol. 139, no. 6, pp. 1713–1722, 2016.
- [38] H. Khan, L. Marcuse, M. Fields, K. Swann, and B. Yener, “Focal onset seizure prediction using convolutional networks,” *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 9, pp. 2109–2118, 2018.
- [39] Y. Li, Y. Liu, W.-G. Cui, Y.-Z. Guo, H. Huang, and Z.-Y. Hu, “Epileptic seizure detection in EEG signals using a unified temporal–spectral squeeze-and-excitation network,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 4, pp. 782–794, 2020.
- [40] H. Peng, C. Lei, S. Zheng et al., “Automatic epileptic seizure detection via Stein kernel-based sparse representation,” *Computers in Biology and Medicine*, vol. 132, Article ID 104338, 2021.
- [41] Z. Deng, P. Xu, L. Xie, K.-S. Choi, and S. Wang, “Transductive joint-knowledge-transfer TSK FS for recognition of epileptic EEG signals,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 8, pp. 1481–1494, 2018.
- [42] M. S. Hossain, S. U. Amin, M. Alsulaiman, and G. Muhammad, “Applying deep learning for epilepsy seizure detection and brain mapping visualization,” *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 15, no. 1s, pp. 1–17, 2019.
- [43] A. Shoeibi, N. Ghassemi, R. Alizadehsani et al., “A comprehensive comparison of handcrafted features and convolutional auto-encoders for epileptic seizures detection in EEG signals,” *Expert Systems with Applications*, vol. 163, Article ID 113788, 2021.
- [44] M. Li and W. Chen, “FFT-based deep feature learning method for EEG classification,” *Biomedical Signal Processing and Control*, vol. 66, Article ID 102492, 2021.
- [45] L. Jiang, J. He, H. Pan, D. Wu, T. Jiang, and J. Liu, “Seizure detection algorithm based on improved functional brain network structure feature extraction,” *Biomedical Signal Processing and Control*, vol. 79, Part 1, Article ID 104053, 2023.
- [46] Y. Zhao, D. Chu, J. He et al., “Interactive local and global feature coupling for EEG-based epileptic seizure detection,” *Biomedical Signal Processing and Control*, vol. 81, Article ID 104441, 2023.
- [47] D. K. Thara, B. G. PremaSudha, and F. Xiong, “Auto-detection of epileptic seizure events using deep neural network with different feature scaling techniques,” *Pattern Recognition Letters*, vol. 128, pp. 544–550, 2019.
- [48] Y. Xiong, J. Li, D. Wu et al., “Seizure detection algorithm based on fusion of spatio-temporal network constructed with dispersion index,” *Biomedical Signal Processing and Control*, vol. 79, Part 2, Article ID 104155, 2023.
- [49] K. Akyol, “Stacking ensemble based deep neural networks modeling for effective epileptic seizure detection,” *Expert Systems with Applications*, vol. 148, Article ID 113239, 2020.
- [50] Y. Pan, X. Zhou, F. Dong, J. Wu, Y. Xu, and S. Zheng, “Epileptic seizure detection with hybrid time-frequency EEG input: a deep learning approach,” *Computational and Mathematical Methods in Medicine*, vol. 2022, Article ID 8724536, 14 pages, 2022.