Proceedings of CONF-CDS 2025 Symposium: Data Visualization Methods for Evaluation
DOI: 10.54254/2755-2721/2025.PO26094

Research on Seizure Detection Using EEG signals Based on Multi-Scale Convolutional Neural Networks BiLSTM-Multi-Head Self-Attention

Tingwei Zhang

School of Computer and Information Technology, Beijing Jiaotong University, Beijing, China
ztw269269@163.com

Abstract. Traditional methods for epileptic seizure detection suffer from limitations such as insufficient feature extraction capability, high computational complexity, and inadequate generalization performance. In this study, leveraging Electroencephalogram (EEG) signals, a novel epileptic seizure detection method based on the Multi-Scale Convolutional Neural Networks-BiLSTM-Multi-Head self-attention(MSCNN-BiLSTM-MHSA) model is proposed. The MSCNN-BiLSTM-MHSA model comprehensively extracts features of EEG signals at different scales by constructing an improved multi-scale convolutional neural network. Furthermore, the introduction of the multi-head self-attention mechanism enables the model to focus more on key features during the iteration process, thereby enhancing the feature learning capability. Experiments were conducted on two epileptic datasets, namely CHB-MIT and Bonn, for validation. The results demonstrate that the MSCNN-BiLSTM-MHSA model achieves an accuracy of 96.23% and 96.13% respectively in epileptic seizure detection.

Keywords: Epileptic Seizure Detection, MSCNN-BiLSTM-MHSA, EEG Signals, Multi-Head Self-Attention Mechanism, Neural Network.

1. Introduction

Epilepsy is a chronic neurological disorder characterized by abnormal electrical discharges in brain neurons. With over 50 million global cases reported by the World Health Organization, it ranks among the most prevalent neurological disorders worldwide[1]. Thus, enhancing the efficiency and accuracy of epileptic seizure detection is critical in neuro-medical research.

Electroencephalogram (EEG) signals, which directly record brain electrical activity with millisecond temporal resolution, are widely used in epileptic detection. They accurately capture epileptiform discharges and dynamically monitor seizure processes, thereby improving detection accuracy. Traditional methods—including SampEn(Sample Entropy), Permutation Entropy, and brain network analysis—have been applied in EEG-based detection (e.g., Zhang et al. combined SampEn with wavelet packet energy features[2]). However, SampEn exhibits limited sensitivity to subtle signal changes, reducing feature extraction accuracy in complex epileptic EEG environments.

To overcome these limitations, deep learning models—such as Convolutional Neural Networks (CNNs)[3], Long Short-Term Memory (LSTM) networks[4], and their variants—have been increasingly adopted. For instance, Ibrahim et al. proposed a CNN-based method, where CNNs

DOI: 10.54254/2755-2721/2025.PO26094

automatically extract spatial features of EEG signals via convolutional and pooling layers, effectively capturing epileptic characteristics[5]. Nevertheless, CNNs perform poorly in processing long-sequence EEG signals due to their weak capability in mining temporal information, limiting their utilization of time-varying features and analysis of dynamic seizure processes.

To address these issues, this paper proposes a multi-scale convolutional attention recurrent neural network (multi-scale CNN-BiLSTM-multi-head self-attention, MSCNN-BiLSTM-MHSA). Firstly, an improved multi-scale CNN enables comprehensive feature extraction, compensating for the inability of traditional fixed-receptive-field networks to capture features of different frequency components. Secondly, a multi-head self-attention mechanism is integrated with a two-layer BiLSTM to better capture temporal dependencies in EEG signals. Finally, the Maximum Kernel Mean Discrepancy method is adopted to enhance the network's feature expression capability.

2. Method and Result

2.1. Data Preparation

CHB-MIT Dataset, sourced from Boston Children's Hospital, serves as a pivotal resource for epilepsy research, primarily encompassing scalp Electroencephalogram (EEG) recordings of pediatric patients with refractory epilepsy[6]. During data acquisition, patients were continuously monitored for several days after discontinuing antiepileptic drugs, aiming to analyze the characteristics of their epileptic seizures and evaluate the feasibility of surgical intervention.

It includes 23 records from 22 subjects (5 males aged 3–22 years and 17 females aged 1.5–19 years), with some records being from the same subject at different time points. These records were acquired at a 16-bit resolution and a sampling frequency of 256 Hz, utilizing the international 10-20 standard electrode positions, and most files contain 23 EEG signals. The total duration of the entire dataset is approximately 967.85 hours, which includes178 recorded epileptic seizures. Data acquisition address: https://archive.physionet.org/pn6/chbmit/.

2.2. Network Model

2.2.1. Convolutional Neural Network

Convolutional Neural Network (CNN) achieve feature extraction of input data through a combination of multiple convolutional layers, pooling layers, fully connected layers, and an output layer. The deep structure formed by the alternating arrangement of convolutional layers and pooling layers enables the hierarchical extraction of deep-seated features from input data[7]. A schematic diagram of the CNN structure is presented in Figure 1.

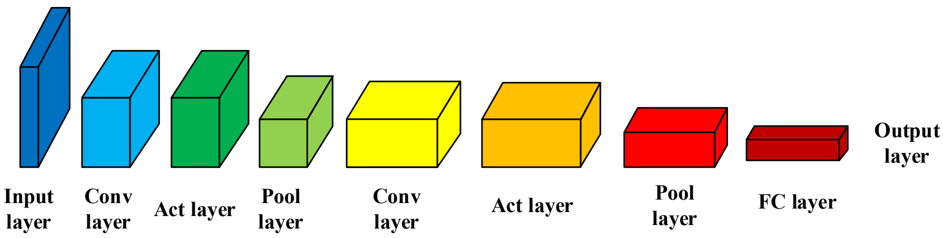


Figure 1. The Structure of The CNN

In convolutional layers, convolutional kernels are used to extract local region features of the input signal, and the feature tensor is obtained from a nonlinear activation function. After performing convolution operations, each convolution kernel will output a feature map[8]. The expression form of convolution operation is equation (1):

DOI: 10.54254/2755-2721/2025.PO26094

$$y^{l(i,j)} = K_i^l * x^{l(j)} = \sum_{j'}^{W-1} K_i^{l(j')} x^{l(j+j')} \tag{1}$$

In equation (1): $K_i^{l(j')}$ is the l weight of the j' convolution kernel in the i layer; The $x^{l(r^j)}$ convolved local region in the l layer; W is the width of the convolution kernel; * for convolution.

After the convolutional layer, a pooling layer is required to reduce the size of the input data, thereby reducing computational complexity and model complexity[9]. The common types of pooling operations are max pooling and average pooling, expressed as follows:

$$p^{l(i,j)} = \frac{1}{W} \sum_{t=(j-1)W+1}^{jW} a^{l(i,t)} \tag{2}$$

$$\gamma^{l(i,j)} = max_{(j-1)W+1 \le t \le jW} \{a^{l(i,t)}\} \tag{3}$$

In equation (2) and (3), $a^{l(i,t)}$ represents the activation value of the t neuron in the i frame of the l layer; W is the width of the pooling area; The activation value of the t neuron in the i frame of the l layer.

2.2.2. Long Short Term Memory

LSTM is an improved recurrent neural network (RNN) model. Effectively solves the problems of gradient vanishing and exploding in traditional RNNs, and enhances the modeling ability of RNNs for long-term dependencies. LSTM effectively prolongs the retention time of past information and reduces the disappearance or explosion of gradients over time by introducing a memory unit called cellular state. In LSTM, cell states can selectively forget or retain information through gating mechanisms[10].

$$f_t = \sigma(W_f * [h_{t-1}, x_t] + b_f) \tag{4}$$

$$i_t = \sigma\left(W_i * \left[h_{t-1}, x_t\right] + b_i\right) \tag{5}$$

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

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{7}$$

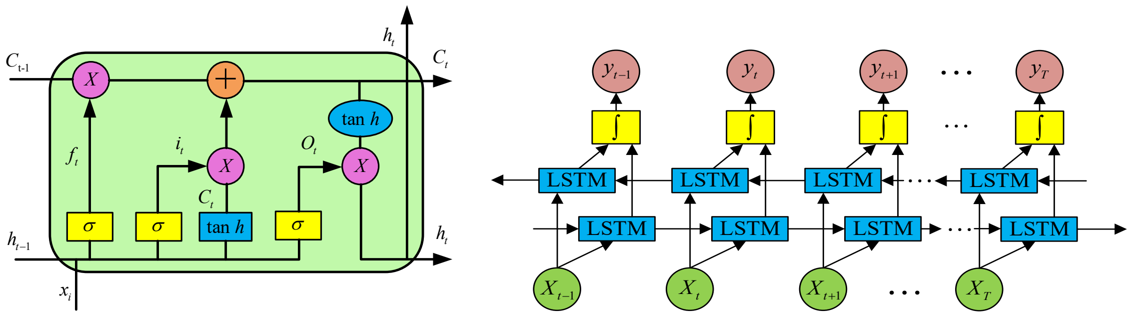
$$O_t = \sigma\left(W_o * \left[h_{t-1}, x_t\right] + b_o\right) \tag{8}$$

$$h_t = O_t * tanh(C_t) \tag{9}$$

In equation (4) - (9): b is the deviation; W is a mixed weight of input and hidden states; σ is the sigmoid activation function. In the t update step, the input gate I, forget gate f output gate O, and cell state c are updated by the input x and hidden state from step n-1. The LSTM model structure is shown in Figure 2 (a).

BiLSTM introduces an additional reverse LSTM layer, allowing the model to simultaneously extract features from the information before and after the current time step. Using different time scales or window sizes for each layer to aggregate features can improve performance and performance in multi-scale modeling tasks. The structure of the BiLSTM model is shown in Figure 2 (b).

DOI: 10.54254/2755-2721/2025.PO26094



(a)The Structure of LSTM (b) The Structure of BiLSTM

Figure 2. The Structure of LSTM and BiLSTM

2.2.3. Multi-Head Self-Attention Mechanism

Interictal Epileptiform Discharges (IEDs) in epileptic EEG signals possess multi-scale characteristics, such as the coupling between high-frequency spikes and low-frequency slow waves, and the contribution of signals from different time periods to detection differs significantly. In this study, a multi-head self-attention module is introduced to dynamically calculate the correlation between features of different time sequences, assigning higher weights to key signal segments. This strengthens the model's ability to capture IED features while suppressing noise interference[11].

The core idea of the self-attention mechanism is to compute the correlation between elements in an input sequence or set and assign weights to each element[12]. These weights reflect the importance of elements in the current task, thus affecting the results of subsequent calculations. A schematic diagram of the self-attention mechanism is shown in Figure 3.

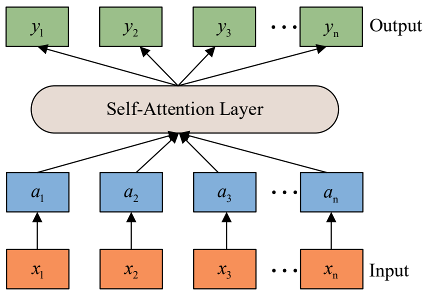


Figure 3. The Structure of IEDs

In traditional self-attention mechanisms, only one attention head is responsible for calculating the correlation score between each position in the input sequence and other positions, and performing weighted summation. The multi head self-attention mechanism introduces multiple different attention heads, each of which can learn different attention representations[13]. This can improve the expression and generalization ability of the model, thereby better addressing complex problems in fault diagnosis. The specific formula are as follows:

$$m - hcad(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = \text{concat}(hcad_1, \cdots, hcad_h)\boldsymbol{w}^0 \tag{10}$$

$$\text{where } hcad_i = \text{Attention}(\mathbf{QW}_i^Q, \boldsymbol{KW}_i^K, \boldsymbol{VW}_i^V) \tag{11}$$

$$\text{attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\boldsymbol{QK}^T}{d^K}\right)\boldsymbol{V} \tag{12}$$

DOI: 10.54254/2755-2721/2025.PO26094

In equation (10) to (12): the Q, K, V are respectively query vector, key vector, and value vector; $\boldsymbol{K}^T$ is the demons of vector K; $\boldsymbol{W}_i^Q \in \Re^{h_{\text{model}} \times d_k}$; $\boldsymbol{W}_i^K \in \Re^{h_{\text{model}} \times d_k}$; $\boldsymbol{W}_i^V \in \Re^{h_{\text{model}} \times d_v}$; $\boldsymbol{W}^O \in \Re^{h d_v \times d_{\text{model}}}$; By incorporating self-attention modules into skip connections, it is hoped that the encoder can preserve the features of normal regions while ignoring the features of defective regions when transmitting information to the decoder, in order to suppress the reconstruction of abnormal regions. The multi-head self-attention mechanism is shown in Figure 4.

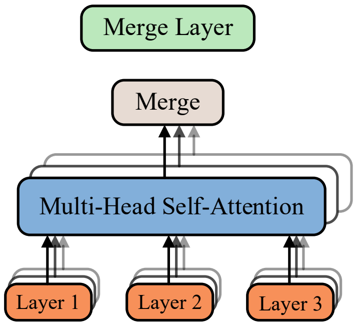


Figure 4. Schematic diagram of multi-head self-attention network structure

2.3. Results and Analysis

When processing the CHB-MIT dataset, bandpass filtering is first used to remove 50Hz power frequency interference, as well as other low-frequency noise and high-frequency interference, limiting the signal frequency range to 0.5Hz-45Hz. This range can effectively preserve the characteristic frequency bands related to epileptic seizures. The classification results of the CHB-MIT dataset are shown in Table 1.

Table 1. Classification results of CHB-MIT dataset

Evaluation Indicators	Accuracy	Precision	Recall	F1-Score
Result (%)	96.23	95.23	96.98	96.10

Table 2. Comparison with other methods

Paper	Method	Accuracy
[14]	Fourier Transform and Pattern Recognition Network	92.5%
[15]	Deep Convolutional Neural Network	90%
[16]	Neural Networks and Wavelet Transform	95.5%
[17]	Artificial Neural Network Based on FPGA	95.14%
This Study	MSCNN BiLSTM MHSA model	96.23%

The proposed multi-scale convolutional attention recurrent neural network performs excellently on the CHB-MIT dataset (Table 1), with 96.23% accuracy for distinguishing epileptic and non-epileptic EEG signals. Its precision (95.23%) and recall (96.98%) reflect advantages in reducing false positives and negatives, with high recall critical for seizure warning; the F1 score (96.10%) verifies its balance.

The MSCNN-BiLSTM-MHSA model was compared with recent epilepsy detection models (Table 2). As shown in Table 3, it achieved 96.23% accuracy in epilepsy detection, outperforming mainstream methods: it outperforms traditional methods like radial basis function neural networks (95.2%) and Fourier transform combined with pattern recognition networks (92.5%) in classification accuracy via model improvements, and surpasses deep convolutional neural networks (90%) and

DOI: 10.54254/2755-2721/2025.PO26094

wavelet transform combined neural networks (95.5%) by enhancing long-sequence modeling while retaining local features.

3. Conclusion

This study proposes a seizure detection method based on MSCNN BiLSTM MHSA. Capturing the multi frequency features of EEG signals through multi-scale convolution, utilizing multi head self attention mechanism to focus on key temporal information, combining two-layer BiLSTM to enhance long-range dependency modeling, and optimizing domain adaptation capability through MK-MMD. On the CHB-MIT dataset, the model achieved an accuracy of 96.23% and an F1 score of 96.10%, while maintaining similar computational efficiency. The ablation experiment verified the necessity of each component and provided an efficient and reliable solution for epilepsy detection.

References

[1] Singh G, Sander J W. The global burden of epilepsy report: Implications for low-and middle-income countries[J]. Epilepsy & Behavior, 2020, 105: 106949.
[2] Zhang, J., Jiang, W., & Ben, X. (2016). Automatic Epileptic Electroencephalogram Detection during Normal, Interictal and Ictal Periods Combining Feature Extraction Based on Sample Entropy and Wavelet Packet Energy with Real AdaBoost Algorithm. Sheng wu yi xue Gong Cheng xue za zhi= Journal of Biomedical Engineering= Shengwu Yixue Gongchengxue Zazhi, 33(6), 1031-1038.
[3] Shoka, A. A. E., Dessouky, M. M., El-Sayed, A., & Hemdan, E. E. D. (2023). An efficient CNN based epileptic seizures detection framework using encrypted EEG signals for secure telemedicine applications. Alexandria Engineering Journal, 65, 399-412.
[4] Shekokar, K. S., & Dour, S. (2022). Automatic epileptic seizure detection using LSTM networks. World Journal of Engineering, 19(2), 224-229.
[5] Ibrahim, F. E., Emara, H. M., El-Shafai, W., Elwekeil, M., Rihan, M., Eldokany, I. M., ... & Abd El-Samie, F. E. (2022). Deep-learning-based seizure detection and prediction from electroencephalography signals. International Journal for Numerical Methods in Biomedical Engineering, 38(6), e3573.
[6] Prasanna, J., Subathra, M. S. P., Mohammed, M. A., Damaševičius, R., Sairamya, N. J., & George, S. T. (2021). Automated epileptic seizure detection in pediatric subjects of CHB-MIT EEG database—a survey. Journal of Personalized Medicine, 11(10), 1028.
[7] Duan, L., Wang, Z., Qiao, Y., Wang, Y., Huang, Z., & Zhang, B. (2021). An automatic method for epileptic seizure detection based on deep metric learning. IEEE Journal of Biomedical and Health Informatics, 26(5), 2147-2157.
[8] Jogin, M., Madhulika, M. S., Divya, G. D., Meghana, R. K., & Apoorva, S. (2018, May). Feature extraction using convolution neural networks (CNN) and deep learning. In 2018 3rd IEEE international conference on recent trends in electronics, information & communication technology (RTEICT) (pp. 2319-2323). IEEE.
[9] Gao Y, Gao B, Chen Q, et al. Deep convolutional neural network-based epileptic electroencephalogram (EEG) signal classification[J]. Frontiers in neurology, 2020, 11: 375.
[10] Sun, M., Song, Z., Jiang, X., Pan, J., & Pang, Y. (2017). Learning pooling for convolutional neural network. Neurocomputing, 224, 96-104.
[11] Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. Neural computation, 31(7), 1235-1270.
[12] Dutta, A. K., Raparthi, M., Alsaadi, M., Bhatt, M. W., Dodda, S. B., Sandhu, M., & Patni, J. C. (2024). Deep learning-based multi-head self-attention model for human epilepsy identification from EEG signal for biomedical traits. Multimedia Tools and Applications, 83(33), 80201-80223.
[13] Xiao, T., Wang, Z., Zhang, Y., Wang, S., Feng, H., & Zhao, Y. (2024). Self-supervised learning with attention mechanism for EEG-based seizure detection. Biomedical Signal Processing and Control, 87, 105464.
[14] Wang, Y., Yang, G., Li, S., Li, Y., He, L., & Liu, D. (2023). Arrhythmia classification algorithm based on multi-head self-attention mechanism. Biomedical Signal Processing and Control, 79, 104206.
[15] Gao Q, Omran A H, Baghersad Y, et al. Electroencephalogram signal classification based on Fourier transform and Pattern Recognition Network for epilepsy diagnosis[J]. Engineering Applications of Artificial Intelligence, 2023, 123: 106479.
[16] Yousefi M R, Dehghani A, Golnejad S, et al. Comparing EEG-based epilepsy diagnosis using neural networks and wavelet transform[J]. Applied Sciences, 2023, 13(18): 10412.

DOI: 10.54254/2755-2721/2025.PO26094

[17] Sarić R, Jokić D, Beganović N, et al. FPGA-based real-time epileptic seizure classification using Artificial Neural Network[J]. Biomedical Signal Processing and Control, 2020, 62: 102106.