# SlimSeiz: Efficient Channel-Adaptive Seizure Prediction Using a Mamba-Enhanced Network

Guorui Lu, Jing Peng, Bingyuan Huang, Chang Gao, Todor Stefanov, Yong Hao, Qinyu Chen

*Abstract*—Epileptic seizures cause abnormal brain activity, and their unpredictability can lead to accidents, underscoring the need for long-term seizure prediction. Although seizures can be predicted by analyzing electroencephalogram (EEG) signals, existing methods often require too many channels or larger models, limiting mobile usability. This paper introduces a SlimSeiz framework that utilizes adaptive channel selection with a lightweight neural network model. SlimSeiz operates in two states: the first stage selects the optimal channel set for seizure prediction using machine learning algorithms, and the second stage employs a lightweight neural network based on convolution and Mamba for prediction. On the Children's Hospital Boston-MIT (CHB-MIT) EEG dataset, SlimSeiz can reduce channels from 22 to 8 while claiming a satisfactory result of 94.8% accuracy, 95.5% sensitivity, and 94.0% specificity with only 21.2 K model parameters, matching or outperforming larger models' performance. We also validate SlimSeiz on a new EEG dataset, SRH-LEI, collected from Shanghai Renji Hospital, demonstrating its effectiveness across different patients. The code and SRH-LEI dataset are available at https://github.com/guoruilu/SlimSeiz.

*Index Terms*—Deep learning, seizure prediction, state-space model, convolutional neural network, healthcare.

## I. INTRODUCTION

Epileptic seizure affects more than 50 million patients around the world according to the World Health Organization (WHO) [1]. Its accompanying symptoms can cause sudden and unforeseen accidents, making patients vulnerable to injury [2]. Therefore, continuous patient monitoring and accurate seizure prediction are essential for improving safety and quality of life. Electroencephalogram (EEG), an electrical recording of brain activity, is a key diagnostic tool for clinicians assessing epilepsy. Fig. 1 illustrates that epilepsy EEG signals can be classified into four states: inter-ictal, pre-ictal, ictal, and post-ictal. Seizure prediction aims at identifying the pre-ictal state. However, due to the complexity of EEG signals and the prolonged nature of seizures, manual identification is impractical, requiring algorithms for automatic identification.

During the past decade, many deep learning techniques have been explored to build seizure prediction systems [3]–[11]. However, most existing methods rely on more than 20 EEG channels to achieve this [3]–[7]. Although effective, using over 20 channels limits patient mobility, increases discomfort, and complicates hardware device design, making these
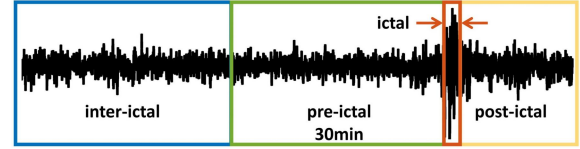
Corresponding authors: Guorui Lu (g.lu@liacs.leidenuniv.nl), Qinyu Chen (q.chen@liacs.leidenuniv.nl) and Hao Yong (haoyong@shsmu.edu.cn).

Guorui Lu, Todor Stefanov, and Qinyu Chen are with the Leiden Institute of Advanced Computer Science (LIACS), Leiden University, The Netherlands.

Jing Peng, Bingyuan Huang, and Hao Yong are with Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China.

Chang Gao is with the Department of Microelectronics, Delft University of Technology, The Netherlands.

Fig. 1. A segment of a patient's scalp EEG from the CHB-MIT dataset. The states can be divided into inter-ictal, pre-ictal, ictal, and post-ictal.

methods unsuitable for mobile and long-term monitoring. Several studies [8]–[11] have explored using fewer channels for seizure prediction. However, some of these works [8], [9] have struggled to achieve 80% prediction accuracy. While other methods [10], [11] have surpassed 90% prediction accuracy using neural network models, they require over 100 K model parameters, making such large models less suitable for long-term use on mobile devices.

To address the aforementioned issues, we propose a framework, called **SlimSeiz**, for efficient seizure prediction using a lightweight Mamba-enhanced neural network with adaptive EEG channel selection. Our main novel contributions are:

1) A machine learning-based adaptive channel selection method that reduces the number of EEG channels from 22 to 8 while maintaining satisfactory accuracy.
2) A lightweight neural network model with only 21.2 K parameters, consisting of one-dimensional convolutions and the Mamba [12] block.
3) A new dataset called the Shanghai Renji Hospital-LEI (SRH-LEI) EEG dataset.
4) An experimental evaluation of SlimSeiz on the Children's Hospital Boston-MIT (CHB-MIT) EEG dataset and the SRH-LEI EEG dataset. SlimSeiz achieves 94.8% accuracy, 95.5% sensitivity, and 94.0% specificity on the CHB-MIT dataset, and 92.7% accuracy, 94.7% sensitivity, and 90.7% specificity on the SRH-LEI dataset.

## II. THE SLIMSEIZ FRAMEWORK

In this section, we start by explaining how the EEG data is pre-processed (Section II-A). Then, we describe the two main parts of SlimSeiz: the channel selection method (Section II-B) and the network model (Section II-C). The channel selection module picks the top $k$ channels that are most important for predictions. The neural network model then makes predictions based on these selected channels.

### A. EEG Data Pre-processing

In this work, we use the CHB-MIT and the newly collected SRH-LEI EEG datasets (details are described in Section III-A).
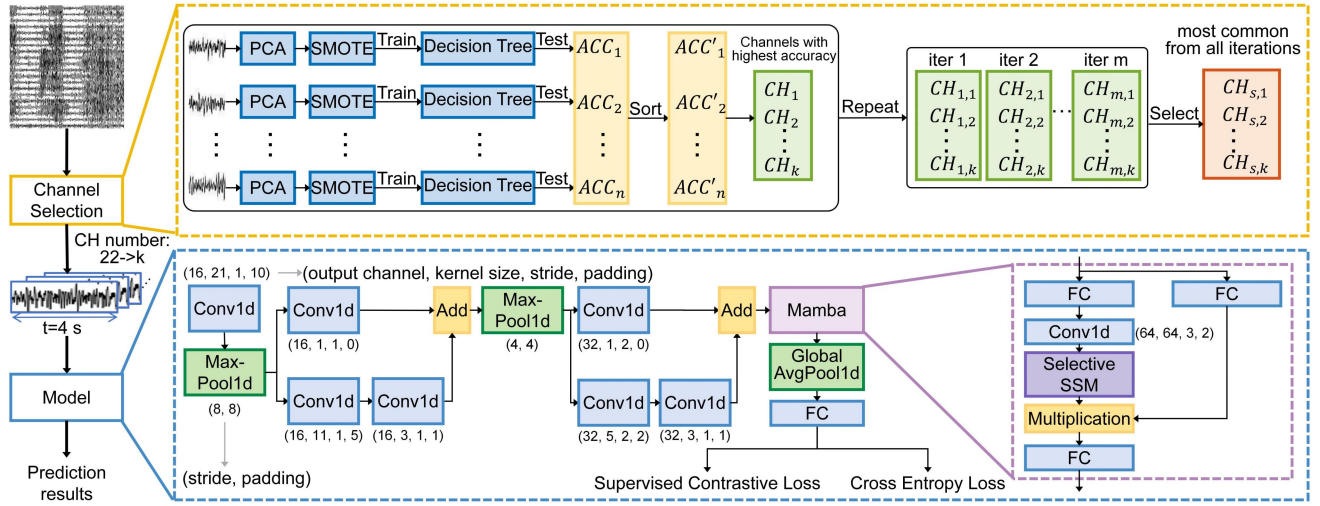
Fig. 2. The overview of SlimSeiz framework. Data from the original 22 channels are input into the channel selection module to select the $k$ (an adjustable hyperparameter) channels that contribute most significantly to seizure prediction, using accuracy (ACC) as the metric. The selected $k$ channels are then sent to the neural network for training and validation. All activation functions are not shown in the figure. The convolutional module in the Mamba block uses the SiLU function, while the other convolution modules use the ReLU function. The FC layers at the input of the Mamba block increase the channel count from 32 to 64, and the output FC layer reduces it from 64 to 32.

Due to the imbalanced nature of the recorded EEG data, where the duration of seizure events is significantly shorter than inter-ictal and post-ictal events, necessary data pre-processing operations can have a considerable impact on the model training and its predictive accuracy. Some previous work [3], [10], [11], [13] discarded the majority of the data to achieve a balanced dataset. In this work, we choose to use overlapping sliding windows with adjustable step sizes to increase the number of pre-ictal segments, similar to the approach used in [5].

The goal of seizure prediction is to identify pre-ictal states. Therefore, ictal segments are discarded because the seizure has already begun. The remaining segments, except pre-ictal, are grouped into a single category for simplicity. We set the pre-ictal segment length to 30 minutes, as prior work indicates this period contains important EEG information for pre-ictal/inter-ictal recognition [14]. If the gap between two seizures is less than 30 minutes, we combine them due to difficulties in distinguishing post-ictal and pre-ictal states. Each EEG recording is split into 4-second segments given as input to the neural network model within SlimSeiz.

### B. Channel Selection Method

Long-term monitoring requires selecting a smaller subset of channels to reduce the computational load on the seizure predicting device and to enhance the patient mobility and comfort. However, the challenge lies in identifying which channels contribute most to the seizure prediction. In our work, we use the prediction accuracy achieved by each individual channel as an indicator of its contribution, based on the assumption that channels with higher prediction accuracy contain more useful information. Although the channel selection is a one-time effort for each patient, we aim to minimize the channel selection time as much as possible. Thus, we propose a channel

selection method based on traditional machine learning that can yield results more quickly compared to neural networks.

The channel selection method is illustrated in the top part of Fig. 2 as a workflow. First, we split each channel data into training and testing segments. Given the fact that traditional machine learning (ML) methods have limited feature extraction capabilities, we use longer 5-second data segments to provide the ML model with more information. Then, we apply Principal Component Analysis (PCA) [15] to reduce the dimension for the training and testing segments separately while preserving key information, thereby reducing subsequent computational overhead. The Synthetic Minority Over-sampling Technique (SMOTE) [16] is employed to balance the dimension-reduced training set, which occupies less memory compared to using overlapping sliding windows. Next, we train and test a decision tree on the processed dataset to obtain the prediction accuracy for the current channel. By executing the aforementioned procedure for each channel, we generate a list of accuracy numbers and select the top $k$ channels based on these accuracy numbers to form a channel subset. To mitigate the influence of random factors, such as the channel data splitting into training and testing segments, on the results, we repeat this procedure $m$ times to obtain $m$ channel subsets.

Finally, we count the occurrences of each channel across all subsets and select the $k$ channels that appear most frequently as the final result. The subset size $k$ and the number of iterations $m$ are two hyperparameters. In this study, we set $m$ to 30 to balance the channel selection time and elimination of random factors. The optimal value of $k$ is determined through relevant experiments to achieve a balance between reducing the number of channels and maintaining high accuracy of the neural network model we have devised to predict seizures.

## C. Neural Network Model Architecture

As shown in the bottom part of Fig. 2, our lightweight model for seizure prediction consists of convolutional feature extraction, Mamba, and classification head.

*1) Convolutional Feature Extraction:* We implement the feature extraction from the EEG input data by stacking one-dimensional (1D) convolutional layers that form the main body of our neural network model for seizure prediction. As shown in Fig. 2, a larger convolution kernel of size 21 is applied on the EEG input data in the first convolutional layer. This is followed by two residual connections, each consisting of three convolutional layers with different smaller kernel sizes. Different kernel sizes help the network extract features at various temporal resolutions. Max pooling and global average pooling are employed to reduce the feature dimensions and computational load. It is worth mentioning that previous work [5], [13], [17]–[19] primarily uses over 20 channels for seizure prediction, treating channels and time as the height and width of an image, respectively, and thus tend to use 2D convolutional neural network to extract spatial and temporal features. In our work, however, we use fewer channels, thereby reducing the complexity of spatial features and making is less necessary to use parameter- and compute-intensive 2D convolutional networks.

*2) Mamba:* The Mamba block is based on the State Space Model (SSM). Its structure, as shown in Fig. 2, uses a fully connected (FC) layer at the input to increase the channel count for capturing more complex features. The subsequent convolutional layer and SSM establish short-term and long-term sequential dependencies, respectively. The FC layer at the output reduces the channel count to maintain consistent input and output dimensions. This structure enables the Mamba block to efficiently capture dependencies within the sequence [20]. Mamba is proposed as an alternative to Transformer to reduce the computational load in the multi-head self-attention mechanism [21] and has shown promising potential in processing physical signals, such as audio signals [22], [23].

Mamba's abilities to capture dependencies in sequential data and reduce computation complexity match our goal to design a lightweight model for processing of EEG signals. To enhance the model's ability to capture temporal dependencies without increasing the model parameters, we add the Mamba block after the convolutional layers. This choice is due to the limited semantic information of sensor signals at a single time point, making it difficult to establish correlations [24]. Therefore, extracting features through convolutional layers and increasing channel count enriches the semantic information at each time point, allowing the Mamba block to work more effectively.

*3) Classification Head:* A fully connected layer is used for classification. Cross-entropy loss and supervised contrastive loss are used to supervise the model's training.

## III. EXPERIMENTAL EVALUATION AND RESULTS

In this section, we evaluate our proposed SlimSeiz framework to demonstrate its potential and efficiency.

TABLE I
DETAILS OF THE SRH-LEI DATASET

| Case No. | Gender | #Seizures | Seizure duration mm:ss | EEG duration hh:mm:ss |
|---|---|---|---|---|
| 1 | Female | 12 | 00:19 | 18:29:23 |
| 2 | Female | 95 | 13:53 | 23:59:50 |
| 3 | Male | 8 | 27:57 | 21:33:56 |
| 4 | Female | 46 | 1:12 | 23:58:51 |
| 5 | Male | 32 | 67:10 | 22:30:45 |
| 6 | Male | 3 | 24:7 | 23:58:51 |
| 7 | Male | 45 | 479:12 | 22:47:10 |
| 8 | Female | 33 | 270:34 | 17:08:34 |

## A. Experimental Setup and Datasets

We evaluate SlimSeiz using two datasets. First, we utilize the CHB-MIT scalp long-term EEG dataset. This dataset comprises recordings from 23 children with intractable seizures at the Children's Hospital Boston, resulting in 24 cases (with two cases derived from a single patient). The duration of EEG recordings for each patient ranges from 19 to 165 hours, totaling approximately 983 hours. The sampling frequency for signal collection is 256 Hz. Second, we introduce and utilize the new SRH-LEI EEG dataset, which we have collected at Shanghai Renji Hospital, Shanghai Jiao Tong University. This dataset includes data from 8 patients, consisting of 4 males and 4 females, aged between 14 and 77 years old. The data collection duration ranges from 17 to 24 hours, with a sampling frequency of 128 Hz. Detailed information about the dateset features is provided in Table I.

We use accuracy (ACC), sensitivity (SENS), and specificity (SPEC) as evaluation metrics. SENS measures the model's ability to correctly identify actual pre-ictal segments, higher sensitivity means fewer missed pre-ictal cases. SPEC indicates how well the model identifies actual inter-ictal segments, higher specificity reflects fewer false alarms for pre-ictal states.

## B. Optimal Number of Channels

To determine the optimal channel counts, we evaluated the model's performance using channel counts of 4, 6, 8, and 10. In this experiment, we randomly split the data into 80% for training and 20% for testing to assess the effectiveness of each channel configuration.

As shown in Fig. 3, our experimental results indicate that increasing the number of channels beyond 8 does not lead to significant performance improvements. More importantly, for most cases (patients), an accuracy of over 90% can be achieved using just 4 channels, and for some patients (e.g., case 01 and 07), the accuracy even exceeds 95%. Therefore, increasing the number of channels offers little benefit for most patients. Only for a few patients (e.g., case 15 and 21) with an accuracy below 90% using 4 channels, increasing the number of channels results in a noticeable performance enhancement (up to 10%). Considering both performance gains and the associated complexity of the device with the number of channels, we conclude that 8 is the optimal number of channels and we use this result in subsequent experiments.
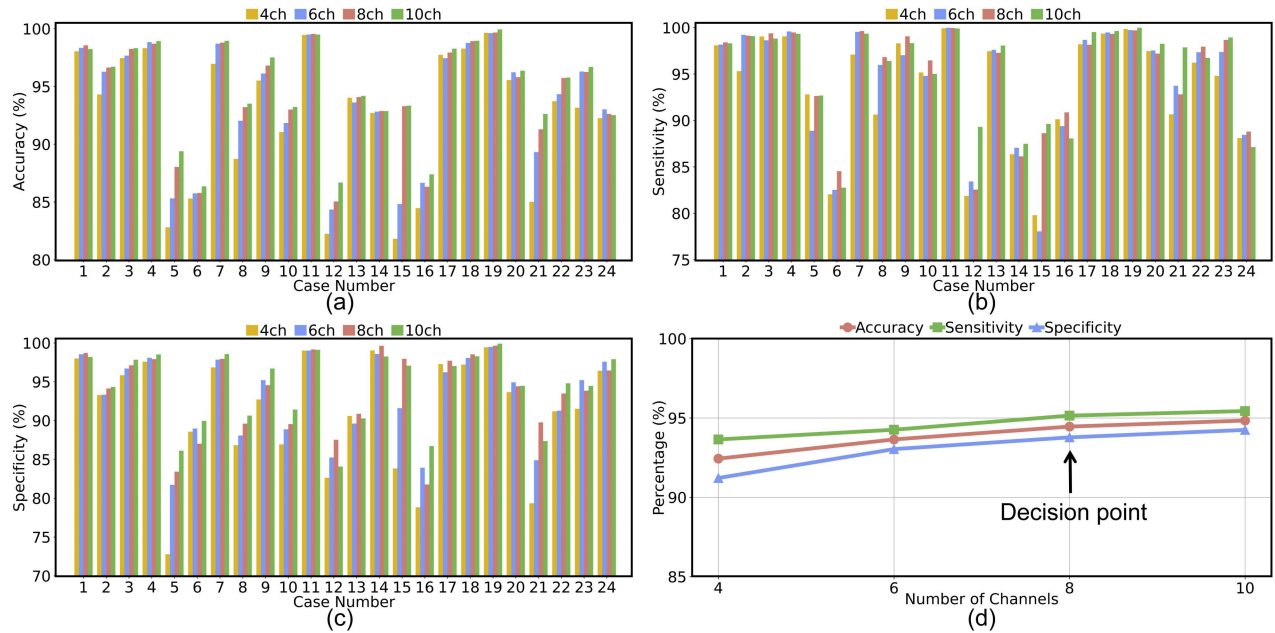
Fig. 3. Performance: (a) accuracy (b) sensitivity (c) specificity across different cases for various channel counts on the CHB-MIT dataset. (d) Overall accuracy, sensitivity, and specificity trends as the number of channels increases from 4 to 10 on the CHB-MIT dataset.

TABLE II
PERFORMANCE COMPARISON OF SLIMSEIZ WITH OTHER WORKS. THE BEST METRIC NUMBER IS **BOLDED**; THE SECOND BEST IS <u>UNDERLINED</u>.

| Ref | Cases | Model | Channels | Parameter | Feature Extraction | Validation | Performance | | |
|-----|-------|-------|----------|-----------|--------------------|------------|-----|------|------|
| | | | | | | | ACC | SENS | SPEC |
| Zhang et all. 2020 [25] | 15 | CNN | <u>18</u> | 194.6K | CSP | LOOCV | 90.0 | 92.2 | 92.0 |
| Buyukccakir et al. 2020 [26] | 10 | MLP | <u>18</u> | 40.9K | HVD | 10-Fold CV | - | 89.9 | - |
| Baghdadi et al. 2020 [27] | 24 | Deep LSTM | <u>18</u> | >3M | Raw EEG | 10-Fold CV | 88.9 | 84.0 | 90.0 |
| Tian et al. 2021 [7] | 7 | Spiking CNN | 23 | **10.3K** | Raw EEG | 80-20 split | - | 95.1 | **99.2** |
| Zhao et al. 2022 [13] | 19 | AddNet-SCL | 22 | 120K | Raw EEG | LOOCV | - | 94.9 | - |
| Lu et al. 2023 [6] | 11 | CBAM-3D CNN-LSTM | 22 | - | STFT | LOOCV | **97.9** | **98.4** | - |
| Chen et al. 2024 [5] | 24 | Spiking Conformer | 22 | 40.3K | Raw EEG | 10-Fold CV | 93.1 | <u>96.8</u> | 89.5 |
| **Our work** | 24 | Convolution Mamba | **8** | <u>21.2K</u> | Raw EEG | 10-Fold CV | <u>94.8</u> | 95.5 | <u>94.0</u> |

TABLE III
PERFORMANCE OF SLIMSEIZ ON THE SRH-LEI DATASET

| Case No. | Accuracy (%) | Sensitivity (%) | Specificity (%) |
|----------|--------------|-----------------|-----------------|
| 1 | 93.7 | 96.1 | 91.2 |
| 2 | 74.3 | 77.6 | 71.0 |
| 3 | 98.0 | 99.0 | 97.0 |
| 4 | 94.5 | 95.3 | 93.7 |
| 5 | 96.0 | 98.8 | 93.3 |
| 6 | 99.0 | 99.5 | 98.4 |
| 7 | 90.5 | 92.3 | 88.6 |
| 8 | 95.6 | 99.1 | 92.2 |
| **Average** | 92.7 | 94.7 | 90.7 |

seizure prediction neural networks.

We also conduct experiments on the SRH-LEI EEG dataset. The results when utilizing 8 selected channels are shown in Table III. SlimSeiz achieves an average of over 90% for each metric, demonstrating its generalization on adult patients.

Finally, we compare our framework with state-of-the-art seizure prediction methods on the CHB-MIT dataset. As shown in Table II, our framework using only 8 channels and a small model (21.2K parameters) achieves comparable average accuracy to methods using over 18 channels and larger models, highlighting its suitability for long-term wearable monitoring.

## C. Performance and Comparison Results

To validate the effectiveness of the Mamba block, we replace it with a convolutional block of a similar parameter scale. The total parameters for the network with the Mamba block and the network with the replaced convolutional block are 21.2K and 22.2K, respectively. While having a slightly smaller model size, the network with the Mamba block achieves an average accuracy that is 0.1% higher than the network with the replaced convolutional block in 10-fold cross validation (CV), indicating the effectiveness of the Mamba block in lightweight

## IV. CONCLUSION AND FUTURE WORK

We present SlimSeiz, a lightweight framework for efficient seizure prediction, using only 21.2 K model parameters and reducing EEG channels from 22 to 8 while matching state-of-the-art performance on the CHB-MIT dataset. Additionally, SlimSeiz is validated on the newly collected SRH-LEI EEG dataset, which will be released soon. We found that Channels P3-O1, P8-O2, C3-P3, C4-P4, FZ-CZ, and P4-O2 are selected in over 17 patients, paving the way for future patient-independent systems with fewer channels.

## REFERENCES

[1] W. H. Organization, *Neurological disorders: public health challenges*. World Health Organization, 2006.

[2] M. Zhou, C. Tian, R. Cao, B. Wang, Y. Niu, T. Hu, H. Guo, and J. Xiang, "Epileptic seizure detection based on eeg signals and cnn," *Frontiers in neuroinformatics*, vol. 12, p. 95, 2018.

[3] A. M. Abdelhameed and M. Bayoumi, "An efficient deep learning system for epileptic seizure prediction," in *2021 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2021, pp. 1–5.

[4] H. Daoud and M. A. Bayoumi, "Efficient epileptic seizure prediction based on deep learning," *IEEE transactions on biomedical circuits and systems*, vol. 13, no. 5, pp. 804–813, 2019.

[5] Q. Chen, C. Sun, C. Gao, and S.-C. Liu, "Epilepsy seizure detection and prediction using an approximate spiking convolutional transformer," *arXiv preprint arXiv:2402.09424*, 2024.

[6] X. Lu, A. Wen, L. Sun, H. Wang, Y. Guo, and Y. Ren, "An epileptic seizure prediction method based on cbam-3d cnn-lstm model," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 11, pp. 417–423, 2023.

[7] F. Tian, J. Yang, S. Zhao, and M. Sawan, "A new neuromorphic computing approach for epileptic seizure prediction," in *2021 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2021, pp. 1–5.

[8] J. S. Ra, T. Li, and Y. Li, "A novel permutation entropy-based eeg channel selection for improving epileptic seizure prediction," *Sensors*, vol. 21, no. 23, p. 7972, 2021.

[9] A. Affes, A. Mdhaffar, C. Triki, M. Jmaiel, and B. Freisleben, "Personalized attention-based eeg channel selection for epileptic seizure prediction," *Expert Systems with Applications*, vol. 206, p. 117733, 2022.

[10] R. Jana and I. Mukherjee, "Deep learning based efficient epileptic seizure prediction with eeg channel optimization," *Biomedical Signal Processing and Control*, vol. 68, p. 102767, 2021.

[11] ——, "Efficient seizure prediction and eeg channel selection based on multi-objective optimization," *IEEE Access*, vol. 11, pp. 54 112–54 121, 2023.

[12] A. Gu and T. Dao, "Mamba: Linear-time sequence modeling with selective state spaces," *arXiv preprint arXiv:2312.00752*, 2023.

[13] Y. Zhao, C. Li, X. Liu, R. Qian, R. Song, and X. Chen, "Patient-specific seizure prediction via adder network and supervised contrastive learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 1536–1547, 2022.

[14] S. Toraman, "Automatic recognition of preictal and interictal eeg signals using 1d-capsule networks," *Computers & Electrical Engineering*, vol. 91, p. 107033, 2021.

[15] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[16] 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.

[17] Y. Song, Q. Zheng, B. Liu, and X. Gao, "Eeg conformer: Convolutional transformer for eeg decoding and visualization," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 710–719, 2022.

[18] T. Dissanayake, T. Fernando, S. Denman, S. Sridharan, and C. Fookes, "Deep learning for patient-independent epileptic seizure prediction using scalp eeg signals," *IEEE Sensors Journal*, vol. 21, no. 7, pp. 9377–9388, 2021.

[19] S. Zhao, J. Yang, and M. Sawan, "Energy-efficient neural network for epileptic seizure prediction," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 1, pp. 401–411, 2021.

[20] H. Qu, L. Ning, R. An, W. Fan, T. Derr, X. Xu, and Q. Li, "A survey of mamba," *arXiv preprint arXiv:2408.01129*, 2024.

[21] X. Zhang, Q. Zhang, H. Liu, T. Xiao, X. Qian, B. Ahmed, E. Ambikairajah, H. Li, and J. Epps, "Mamba in speech: Towards an alternative to self-attention," *arXiv preprint arXiv:2405.12609*, 2024.

[22] Y. Chen, J. Yi, J. Xue, C. Wang, X. Zhang, S. Dong, S. Zeng, J. Tao, L. Zhao, and C. Fan, "Rawbmamba: End-to-end bidirectional state space model for audio deepfake detection," *arXiv preprint arXiv:2406.06086*, 2024.

[23] J. Lin and H. Hu, "Audio mamba: Pretrained audio state space model for audio tagging," *arXiv preprint arXiv:2405.13636*, 2024.

[24] Y. Zhang, L. Wang, H. Chen, A. Tian, S. Zhou, and Y. Guo, "If-convtransformer: A framework for human activity recognition using imu fusion and convtransformer," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 6, no. 2, pp. 1–26, 2022.

[25] Y. Zhang, Y. Guo, P. Yang, W. Chen, and B. Lo, "Epilepsy seizure prediction on eeg using common spatial pattern and convolutional neural network," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 2, pp. 465–474, 2019.

[26] B. Büyükçakır, F. Elmaz, and A. Y. Mutlu, "Hilbert vibration decomposition-based epileptic seizure prediction with neural network," *Computers in biology and medicine*, vol. 119, p. 103665, 2020.

[27] A. Baghdadi, R. Fourati, Y. Aribi, P. Siarry, and A. M. Alimi, "Robust feature learning method for epileptic seizures prediction based on long-term eeg signals," in *2020 international joint conference on neural networks (IJCNN)*. IEEE, 2020, pp. 1–7.