A Lightweight End-to-End Neural Networks for Decoding of Motor Imagery Brain Signal

Hyeon Kyu Lee, Ji-Soo Myoung, Young-Seok Choi
Department of Electronics and Communications Engineering
Kwangwoon University
Seoul, Republic of Korea
{skgusrb12, qndjd206, yschoi}@kw.ac.kr

***Abstract*—In motor imagery-based brain-computer interface (MI-BCI), the variants of convolutional neural networks (CNNs) have been increasingly received attention due to relatively outstanding decoding performance. However, the growing network size for high decoding performance and the inefficient procedures of BCI systems lead to limited availability in real-life MI-BCI systems. To tackle these issues, we propose an end-to-end neural network named lightweight EEG-inception squeeze-and-excitation network (LiteEEG-ISENet). The architecture is built to remedy the two parts: 1) depthwise convolution is adopted to reduce the computational complexity of the network and train the intrinsic features for each channel of the MI dataset; 2) In addition to the previous motivation, the squeeze-and-excitation (SE) blocks are employed to recalibrate channel-wise feature response adaptively. The experimental results on the public dataset widely used in the MI-BCI study demonstrate that the proposed method outperforms the existing method in terms of decoding performance and neural network memory efficiency.

Keywords—Brain-computer interface, motor imagery, electroencephalography, neural network, brain signals

I. Introduction

A motor imagery-based brain-computer interface (MI-BCI) is a system that communicates and controls various surrounding devices by decoding electroencephalography (EEG) signals with the implicit intention of the user. In such MI-BCI systems, it has been an open challenge to decode the user's intention in that EEG signals possess statistical complexity and are easily distorted by artifacts, movements, eye blinks, and so on [1]. Recently, numerous neural networks motivated by the original convolutional neural networks (CNNs) have shown record-breaking performances compared with the existing methods [2-4].

In general, the realization of BCI systems consists of signal acquisition, preprocessing, feature extraction, classification, and transferring the output of classification into control command. However, it is time-consuming and laborious to choose suitable feature extraction and classification methods for high classification accuracy in a dynamic background environment. Due to the recent advent of deep learning, several end-to-end CNN-based neural networks with high performance and simple architecture have been introduced [5, 6]. Recently, Zhang *et al*. presented the CNN model for MI-BCI based on an inception-time network with an efficient and accurate time-series decoding capacity, named an EEG-inception [7]. This network has been designed to capture high-quality time-series features by deep blocks and multi-branch inception structures corresponding to multiple time windows. However, large-scale neural networks, including EEG-inception, are unavailable for real-life MI-BCI systems due to limited computing resources.

In this study, we propose a lightweight end-to-end neural network while improving decoding performance, which is called a lightweight EEG-inception squeeze-and-excitation network (LiteEEG-ISENet). We focus on two issues: (1) designing a memory-efficient neural network architecture in a resource-constrained environment (2) considering the relatively low-performance problem in the light model. To deal with the former issue, we adopt the depthwise convolution that saves computation by operating for each channel of input data [8]. In addition, from a neurophysiological point of view, some channels have a more critical effect on MI tasks [9]. In the latter issue, to further reinforce channel-wise dependencies, the attention mechanism of squeeze-and-excitation (SE) blocks is utilized, which contributes to improving the classification accuracy by capturing the channel-wise features related to MI tasks [10]. The proposed method was evaluated by the BCI competition IV dataset 2a widely used in MI-BCI studies. Based on the experiment results, LiteEEG-ISENet demonstrates its superiority by showing high classification accuracy in most subjects despite about 10× fewer parameters.

II. Methods

In this section, we describe the proposed method and background. First, we introduce depthwise convolution and SE blocks used in this study. Then, we illustrate the overall scheme of the proposed model base on the previous methods, which is summarized in Figure 1.

A. Depthwise Convolution

The depthwise convolution approach is the most often used method for reducing the number of parameters and processing cost of a normal convolution-based neural network. Here the convolution operations in each channel of input data are separately performed. Therefore, the number of parameters in the model can be diminished from d^2k to dk where k is the kernel width. The output of the depthwise convolution is given by [8, 11]:

$$\text{DepthwiseConv}(X, w_{c,:}, i, c) = \sum_{j=1}^{k} w_{c,j} \cdot X_{\left(i+j-\left[\frac{k+1}{2}\right]\right),c} \quad (1)$$

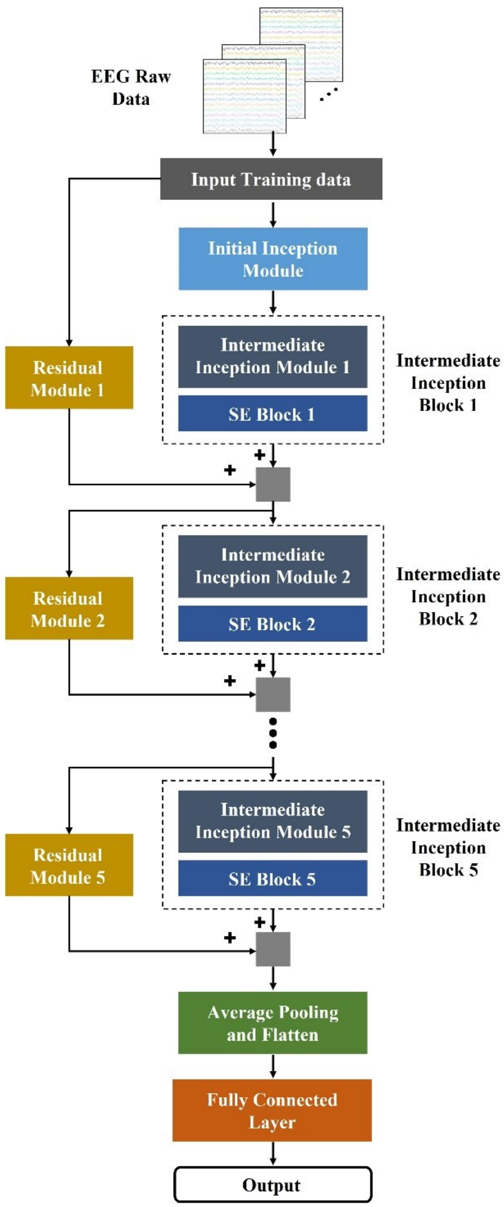

Fig. 1. Overall visualization of the LiteEEG-ISENet architecture

where $w \in \mathbb{R}^{d \times k}$ indicate weight for element i and c is output dimension.

B. Squeeze-and-Excitation Network (SENet)

According to [10], the squeeze-and-excitation network (SENet) is comprised of SE blocks that adaptively recalibrate a channel relationship by squeezing the features using global average pooling, which has shown the compelling performance for the ImageNet database. SENet consists of two steps, squeeze and excitation. In the first step, all features along the channel axis are converted into a one-dimensional vector with condensed channel-wise representations. Then, the multi fully-connected layers in the excitation step can be obtained channel-wise

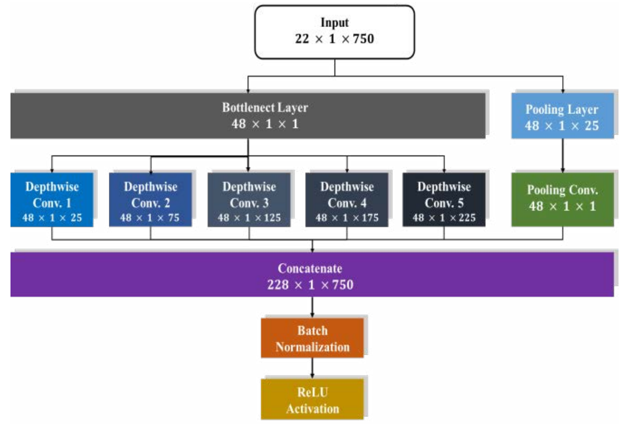

Fig. 2. Overall architecture of the proposed initial module and intermediate module in LiteEEG-ISENet. Here, the initial and intermediate inception modules have the input data depth of 22 and 288, respectively.

attention from the output of the squeeze step. A more detailed description is provided in [10].

C. A Lightweight EEG-Inception Squeeze-and-Excitation Network (LiteEEG-ISENet)

We propose a light and more robust end-to-end neural network by introducing depthwise convolution and SE block into the cornerstone network of our study, EEG-inception [7]. As shown in Fig. 1., the proposed LiteEEG-ISENet consists of one initial inception module, five intermediate modules, five SE blocks, five residual modules. Each inception module includes a bottleneck layer, multiple 1D convolutional layers, a 1D max-pooling layer, a batch normalization, and a ReLU activation function. The overall scheme is employed in the identical structure as in [7]. The bottleneck layer with [1× 1] kernel matrix expands the dimension of the time-series EEG input data from 22 to 48. Then, to reduce the computational complexity of the network, we replace all standard 1D convolutional layers after the bottleneck layer with the 1D depthwise convolution layers, and the detailed architecture is illustrated in Fig. 2. The five depthwise convolution operations include the kernel sizes of [25 × 1], [75 × 1], [125 × 1], [175 × 1], and [225 × 1], respectively. Each kernel size indicates the size of the window along the time axis, i.e., 0.1 s, 0.3 s, 0.5 s, 0.7 s, and 0.9 s. The pooling layer is applied to collect the various feature by downsampling for features. The different features extracted by multiple branches are concatenated, which are followed by batch normalization and ReLU activation function.

The small-scale neural network generally tends to reveal relatively low performance than the large-scale neural network. We attach the SE block after every intermediate module to complement this fault. In the SE block, in turn, channel-wise attention is obtained, and the channel reduction ratio is set to 8. Then, the outputs of the SE block and intermediate module are calculated by channel-wise multiplication.

The residual module is employed to tackle the learning degradation problem generated in deep neural networks. In addition, this module consists of a convolutional layer, a batch normalization, and a ReLU activation function. It is applied after every intermediate inception block.

TABLE I. COMPARISON OF THREE NEURAL NETWORKS

Networks	Parameters	FLOPs
EEG-inception	8.88M	2.56G
LiteEEG-inception	417.99k	140M
LiteEEG-SE-inception	858.52k	235M

TABLE II. THE SUBJECT-INDEPENDENT CLASSIFICATION ACCURACY (%) BY DIFFERENT NETWORKS ON BCI COMPETITION IV DATASET 2A

Subjects	Networks (%)		
	EEG-inception	LiteEEG -Inception	LiteEEG -ISENet
S1	75.86	77.58	**79.31**
S2	57.76	**64.65**	57.75
S3	85.34	**92.24**	87.93
S4	56.89	60.34	**62.93**
S5	68.1	65.51	**75.86**
S6	**62.07**	59.48	59.48
S7	87.93	87.93	**91.37**
S8	82.76	86.2	**90.51**
S9	75.86	76.72	**83.62**
Average	72.51	74.51	**76.52**

Finally, the time-wise dimensionality reduction in output of the fifth intermediate block is performed by average pooling. The average pooling is followed by a fully connected layer having four outputs corresponding to four MI tasks.

III. EXPERIMENTAL RESULTS

A. Datasets

To evaluate the proposed method, we have utilized BCI competition IV dataset 2a. The nine subjects participated in the experiment including a total of two sessions with 288 trials per session. Each trial has 22 EEG channels, a sampling rate of 250 Hz, and four MI tasks (left hand, right hand, feet, and tongue). Here, we only utilized MI periods between 3 s and 6 s to classify the four MI tasks.

In the experiment, the evaluation dataset is split into five folds (four for training and one for test). Then, to train the proposed method, the training dataset has performed by 10-fold cross-validation.

B. Evaluation Performance

All the experiments are implemented on Ubuntu 18.04.6 LTS, CPU: AMD Ryzen Threadripper 3960X 24-Core Processor, GPU: Nvidia Geforce RTX 3090, RAM: 64 GB, Python: 3.8.12, PyTorch: 1.10.1.

Table I reports the results of the number of parameters and of the computation for EEG-inception and the proposed methods with the same input size. Here, FLOPs indicate floting points operations. For each result, the proposed LiteEEG-ISENet shows 90.33% and 90.8% less than that of the EEG-inception. Note that the number of parameters and of the computation for proposed LiteEEG-inception is much smaller than that of the other networks, while there is a low decoding performance than LiteEEG-ISENet, which is shown in Table II.

In Table II, the results for subject-independent classification accuracy are summarized. From the table, we can confirm the superiority of the proposed LiteEEG-ISENet in that it results in high classification accuracy across most subjects. Notably, it shows more than enhanced classification accuracy for subjects 5 and 7. Moreover, in terms of the average classification accuracy, LiteEEG-ISENet has achieved a 4.01% higher than the baseline method.

IV. CONCLUSION

In this paper, we present the LiteEEG-ISENet, which is motivated to capture channel-wise dependencies while diminishing the size of the network. The experimental results show the potentiality of the proposed method in that it achieves improved classification accuracies across multiple subjects on the public MI evaluation dataset. In addition, the proposed LiteEEG-ISENet reduces the computational cost of the EEG-inception base model by about 10× and does not require inconvenient procedures such as preprocessing and feature extraction. Consequently, the proposed end-to-end neural network may provide a more suitable tool for real-life MI-BCI systems.

ACKNOWLEDGMENT

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2021-0-00034, Clustering technologies of fragmented data for time-based data analysis).

REFERENCES

[1] F. Lotte *et al.*, "A review of classification algorithms for EEG-based brain– computer interfaces: A 10 year update," *J. Neural Eng.*, vol. 15, no. 3, Apr. 2018.

[2] H. K. Lee and Y.-S. Choi, "Application of continuous wavelet transform and convolutional neural network in decoding motor imagery brain-Computer Interface," *Entropy,* vol. 21, no. 12, 2019.

[3] X. Xiao and Y. Fang, "Motor imagery EEG signal recognition using deep convolution neural network," *Front Neurosci*. vol. 15. 2021.

[4] B. Sun, X. Zhao, H. Zhang, R. Bai, and T. Li, ''EEG motor imagery classification with sparse spectrotemporal decomposition and deep learning,'' *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 2, pp. 541–551, Apr. 2021.

[5] V. J. Lawhern *et al.*, "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, Oct. 2018.

[6] J. Liu, F. Ye, and H. Xiong, "Multi-class motor imagery EEG classification method with high accuracy and low individual differences based on hybrid neural network," *J. Neural Eng.*, vol. 18, no. 4, Aug. 2021.

[7] C. Zhang, Y.-K. Kim, and A. Eskandarian, "EEG-inception: an accurate and robust end-to-end neural network for EEG-based motor imagery classification," *J. Neural Eng.*, vol. 18, no. 4, Mar. 2021.

[8] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv:1704.04861*, 2017.

[9] H. Zhang, X. Zhao, Z. Wu, B. Sun, and T. Li, "Motor imagery recognition with automatic EEG channel selection and deep learning," *J. Neural Eng.*, vol. 18, no. 1, Feb. 2021.

[10] J. Hu, L. Shen, and G. Sun, ''Squeeze-and-excitation networks,'' in Proc. Conf. Comput. Vis. Pattern Recognit. (CVPR), 2018, pp. 7132–7141.

[11] Wu, A. Fan, A. Baevski, Y. Dauphin, and M. Auli, "Pay less attention with lightweight and dynamic convolutions," in Proc. Int. Conf. Learn. Representations (ICLR), 2019