Undergraduate Conference on Intelligent Computing and Systems (UCICS 2025)
26-27 February, 2025; Varendra University, Rajshahi, Bangladesh

Data Augmentation Approach to Frequency Recognition of SSVEP Using Mask Encoding Combination Based Deep Learning

Md Nayem Molla
Dept. of Computer Science and Eng.
University of Rajshahi, Bangladesh
esnayemmolla@gmail.com

Tareq Munawer Siddiqui
Dept. of Computer Science and Eng.
University of Rajshahi, Bangladesh
tareqmunawertaj@gmail.com

Shakil Hossan
Dept. of Computer Science and Eng.
University of Rajshahi, Bangladesh
shakilhossan113@gmail.com

Sk Solaiman Abdullah
Dept. of Computer Science and Eng.
University of Rajshahi, Bangladesh
solaiman76149@gmail.com

Zannatul Naim
Dept. of Computer Science and Eng.
University of Rajshahi, Bangladesh
zannatulnaim09@gmail.com

Shabbir Mahmood
Dept. of Computer Science and Eng.
PUST, Pabna, Bangladesh
shabbir.cse@pust.ac.bd

***Abstract*—Steady-state visual evoked potential (SSVEP) based brain computer interfaces (BCIs) are promising technique for real time communication and control. Utilizing transfer learning technique, this investigation introduces a novel classification method that incorporates mask encoding combination (MEC) data augmentation and convolutional neural networks (CNNs). The method's superior classification performance is achieved by processing harmonics, channels, and temporal sub-bands, which enhances the robustness of multi-channel EEG signal analysis. In a 1s time window, the approach obtains a maximal accuracy of 94.69% and a peak information transfer rate (ITR) of 193.14 bits min^{-1} when evaluated on a benchmark dataset of 35 subjects and 40 characters. These findings surpass those of conventional methodologies, emphasizing the potential of integrating data augmentation and transfer learning to accelerate the development of SSVEP-based BCIs.**

***Index Terms*—Steady-state visual evoked potential (SSVEP), Brain-computer interface (BCI), Convolutional neural network (CNN), Mask encoding combination (MEC) data augmentation.**

I. Introduction

A brain-computer interface (BCI) establishes communication between the human brain and an external device, directing specific activities without relying on peripheral nerve and muscle actions. This technology involves identifying brain activity through neurophysiological signals, creating corresponding commands for interaction with external devices. This direct communication pathway enhances control over the connected device [1]. The Electroencephalogram (EEG) is a widely used method for capturing brain activity in BCI implementations, known for its noninvasiveness and high time resolution [2].
Numerous methods have been devised to recognize the frequency of SSVEP, with the accuracy of frequency recognition algorithms being crucial and challenging in SSVEP-based BCI development. Recently, deep learning-based approaches have emerged as powerful alternatives, capable of automatically extracting features from raw EEG data without the need for manually designed feature engineering. This work is organized by first introducing EEG signals and BCIs, then concentrating on the difficulties in recognizing SSVEP frequencies. Next, a full description of the suggested deep learning model and EEG mask encoding (EEG-ME) for data augmentation is provided. In order to show how well the method works to improve classification performance, the results are finally assessed.

II. Problem Statement

SSVEP-based BCIs still face significant challenges in accurately identifying targets because of the inherent variability in electroencephalography (EEG) signals and the shortcomings of conventional manually designed feature-based techniques. Although deep learning techniques provide automated feature extraction, they are not sufficiently strong to handle a variety of EEG patterns. In order to increase classification accuracy and robustness in SSVEP-based BCIs for practical applications, this study suggests a unique transfer learning architecture that makes use of CNNs and EEG Mask Encoding (EEG-ME) for data augmentation.

III. Related Works

The creation of brain-computer interfaces, or BCIs, in recent decades has made it possible for the human brain to communicate directly with external devices via neurophysiological signals, obviating the requirement for actions from the muscles and peripheral nerves [3]. The noninvasive nature and affordability of EEG-based BCIs have made them popular in neural engineering, neuroscience, and clinical rehabilitation. They hold great promise for people with severe motor disabilities, including those suffering from spinal cord injuries, locked-in syndrome, and comas [4].
Due to its greater ITR, signal-to-noise ratio (SNR), and low user training needs, the SSVEP-based BCI has gained attention [5]. Conventional techniques for classifying SSVEPs

have been widely employed, such as power spectrum density analysis (PSDA) and canonical correlation analysis (CCA) [1], correlated component analysis (CORRCA) [6]. Nevertheless, it has been demonstrated that using individual-specific data can greatly enhance performance, which current approaches do not achieve. By simultaneously learning temporal and spatial EEG characteristics, deep learning algorithms have recently been used to overcome these restrictions and improve classification accuracy [7]. In practical applications, techniques such as CNNs and transfer-related component analysis (TransRCA) have shown enhanced performance, resolving problems such inter-subject variability and lowering the requirement for large amounts of training data [8], [9]. Further advancements that have showed promise in enhancing SSVEP-based BCI performance include EEG-ME for data augmentation [10].

IV. Dataset and Preprocessing

A. Data-Set Description

The Benchmark SSVEP Dataset, which was developed by the Tsinghua group and is publicly available, was employed in this experiment. It comprises EEG recordings from 35 participants (18 males, 17 females, mean age: 22 years), eight of whom had prior exposure with SSVEP-based BCIs. The experiment consisted of the presentation of 40 visual stimuli that flickered at frequencies ranging from 8 Hz to 15.8 Hz, with intervals of 0.2 Hz. Each participant completed six sessions of 40 trials. A 0.5s visual cue was followed by 5s of simultaneous stimulus flickering and a 0.5s vacant screen in each trial, which lasted 6 s. Participants were able to maintain their concentration by observing a red triangle situated beneath the target stimulus.

B. Data Preprocessing

The continuous EEG data were segmented into 6-second epochs, which captured 0.5 seconds of pre-stimulus and 5.5 seconds of post-stimulus activity. The downsampling of these segments to 250 Hz resulted in 1500 time values per trial. The data of each subject was stored in a distinct file (e.g., S01.mat, S02.mat). This file contained a 4-dimensional matrix that was organized by the electrode index (64 channels), time points (1500 time points per trial), target stimuli (40 stimuli), and block index (6 blocks/trials). In order to guarantee comprehensive data organization for subsequent analysis, electrode positions and information on stimulus frequencies and phases were recorded in distinct files.

C. Data Augmentation

Deep learning (DL)-based methods for EEG classification often struggle with limited data availability in public datasets, leading to overfitting and reduced performance. To overcome this, data augmentation is a useful approach. This study introduces a technique called MEC to enhance DL-based SSVEP classification. MEC involves two main steps: mask encoding and average combination.

Mask Encoding: Mask Encoding is a data augmentation technique used to enhance the robustness of deep learning models by introducing variability in EEG training data.In the following figure X-axis represents the time(t) and Y-axis represents the amplitude(A).

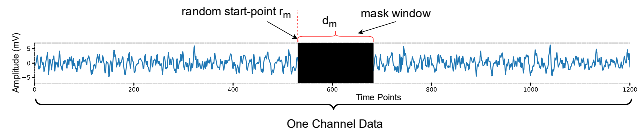

Figure 1. Mask Encoding for One Electrode Channel Data of a Sample

- EEG Data Representation: The EEG signal is represented as a matrix:
$$X \in \mathbb{R}^T$$
where T is the number of time steps (data points).
- Mask Ratio(r_m): The fraction of the signal to be masked, with $r_m \in [0, 1]$, where (r_m) is the mask ratio.
- Length of Mask Window (d_m): d_m is the length of the mask window, defined as:
$$d_m = T \times r_m \tag{1}$$
- Starting Point of Mask (r_s): The starting point of the masked segment is randomly selected as:
$$r_s \in [0, T - d_m]$$
- Masked Data ($X_{\text{masked}}(t)$):
$$X_{\text{masked}}(t) = \begin{cases} 0, & \text{if } r_s \leq t < r_s + d_m, \\ X(t), & \text{otherwise.} \end{cases} \tag{2}$$

Average Combination: After mask encoding, 4 masked trials are generated. In the average combination step, 11 new trials are created by averaging different combinations of these four trials. This increases data variability and improves the model's robustness.

TABLE I
Average Combinations of Masked Trials

Trial	Combination	Description
1	(1, 2)	Average of Masked Trials 1 and 2
2	(1, 3)	Average of Masked Trials 1 and 3
3	(1, 4)	Average of Masked Trials 1 and 4
4	(2, 3)	Average of Masked Trials 2 and 3
5	(2, 4)	Average of Masked Trials 2 and 4
6	(3, 4)	Average of Masked Trials 3 and 4
7	(1, 2, 3)	Average of Masked Trials 1, 2, and 3
8	(1, 2, 4)	Average of Masked Trials 1, 2, and 4
9	(1, 3, 4)	Average of Masked Trials 1, 3, and 4
10	(2, 3, 4)	Average of Masked Trials 2, 3, and 4
11	(1, 2, 3, 4)	Average of Masked Trials 1, 2, 3, and 4

V. Methodology

For better frequency recognition, the suggested methodology for the SSVEP-based BCI system makes use of CNNs and transfer learning. The weights of a broad model that has

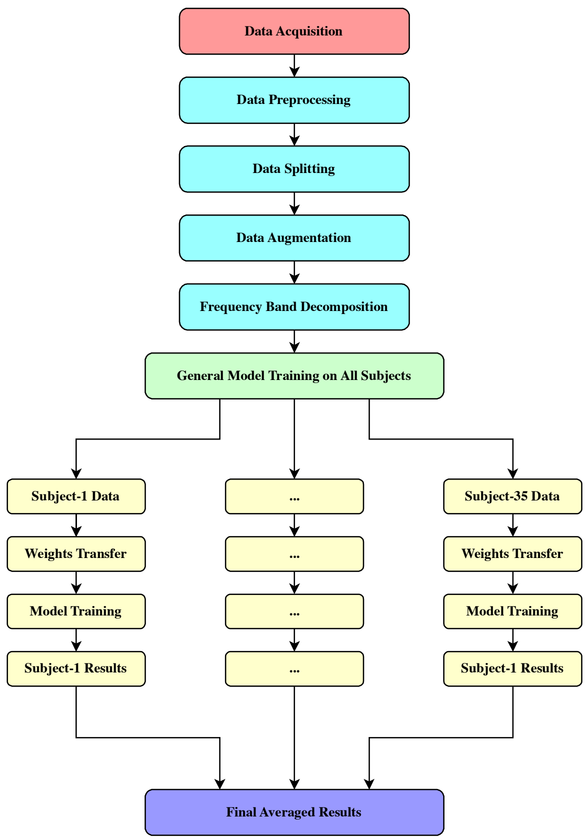

Figure 2. Block Diagram of Proposed Method

been trained on data from every subject are then transferred to subject-specific models for fine-tuning. This individualized method addresses inter-subject variability, improving accuracy and ITR. The robustness and generalization of the model are further strengthened by the application of data augmentation techniques.

A. CNNs for EEG Classification

CNNs are effective in EEG classification by learning spatial and temporal features from EEG data. The input is represented as a 3D tensor $X \in \mathbb{R}^{C \times D \times S}$, where C is the number of channels, D is the number of time points, and S is the number of frequency bands. The convolutional layers extract features using filters, with the output feature map calculated as:

$$Z_{i,j} = \sum_{m=0}^{f_H-1} \sum_{n=0}^{f_W-1} X_{i+m,j+n} \cdot W_{m,n} + b \tag{3}$$

ReLU activation $\text{ReLU}(x) = \max(0, x)$ introduces non-linearity, and pooling layers reduce the spatial dimensions with max pooling

$$P_{i,j} = \max_{(m,n)\in W} Z_{i+m,j+n} \tag{4}$$

Dropout is applied to prevent overfitting, where the Hadamard product is used to apply the dropout operation, resulting in the final dropout output $Z_{\text{drop}} = Z \odot D$ and D is a binary mask. The data is then passed through fully connected layers, with outputs computed as:

$$y_i = w_i^T \cdot f + b_i \tag{5}$$

The output layer uses softmax activation to calculate probabilities:

$$p(y = i|x) = \frac{e^{y_i}}{\sum_{j=1}^{K} e^{y_j}} \tag{6}$$

The model is trained using the categorical cross-entropy loss:

$$L = -\sum_{i=1}^{K} y_i \log(p(y = i|x)) \tag{7}$$

Optimization is done using Adam optimizer algorithm. Once trained, the model predicts the class

$$\hat{y} = \arg\max_i p(y = i|x) \tag{8}$$

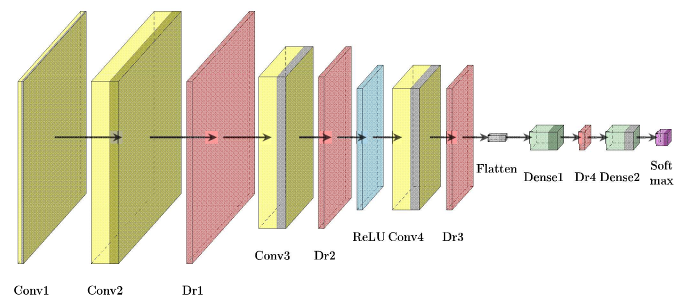

Figure 3. Proposed CNN Architecture

B. Proposed Frequency Recognition Method

This section outlines the methodology used for recognizing SSVEPs frequencies from EEG signals. The process consists of the following key stages, as illustrated in Figure-2:

- Data Preprocessing: Cleaning EEG signals and select relevant channels/time windows for high-quality input.
- Data Partitioning: Split the dataset into training, validation, and test sets.
- Data Augmentation: Apply mask-encoding to diversify training data and improve model generalization.
- Frequency Band Decomposition: Isolate EEG frequency bands for feature extraction.
- Model Training and Transfer Learning: Train a general CNN model on data from all subjects. Transfer learned weights to subject-specific models for fine-tuning.
- Performance Evaluation: Compute classification accuracy for all subjects and average the results to measure overall performance.

VI. Results And Performance Analysis

With an emphasis on two important metrics mean accuracy and ITR, this study contrasts the suggested method with a number of cutting-edge techniques for SSVEP-based BCIs. When assessing the practical usability of BCIs, several parameters are crucial.

A. Experimental Results

TABLE II
EXPERIMENTAL RESULTS WITH TEST ACCURACY AND INFORMATION TRANSFER RATE (ITR) BITS MIN^{-1} (1S)

Experiment	Accuracy (%)	ITR (bits min^{-1})
Experiment 1 (9 EEG channels)	93.23	188.38
Experiment 2 (10 EEG channels with data aug.)	94.69	193.14

This improvement is attributed to the use of additional channels and data augmentation techniques, which enhanced the model's robustness to noise and variability, and its generalization across subjects.

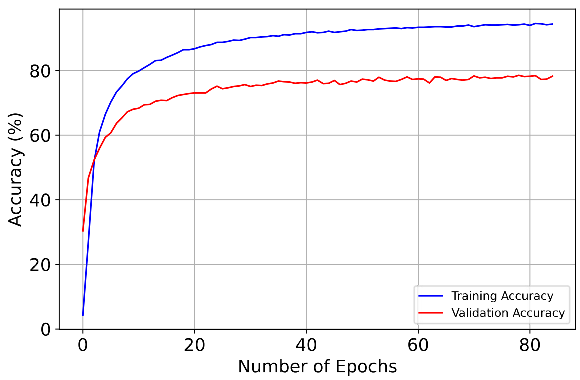

Figure 4. CNN Accuracy Graph

B. Comparison with Other Methods

TABLE III
COMPARISON OF ACCURACY AND ITR FOR DIFFERENT METHODS AT 1S TIME WINDOW

Method (Reference)	Year	Accuracy (%)	ITR (bits min^{-1})
SSVEPformer [9]	2023	83.19	157.65
3DCNN-TL [8]	2024	89.35	173.02
ASS-IISCCA [11]	2023	93.00	175.00
Conv-CA [7]	2020	93.88	190.23
Proposed Method (with 9 channels)	-	93.23	188.38
Proposed Method (10 channels with data aug.)	-	**94.69**	**193.14**

SSVEPformer [9] accuracy was 83.19%, and the ITR was 157.65 bits min^{-1}. Transformer models, while effective for certain applications, require large datasets and computational resources, limiting their real-time applicability in BCIs. Another method 3DCNN-TL [8] accuracy found 89.35% and ITR was 173.02 bits min^{-1}. This model struggles with generalization due to dataset variations. ASS-IISCCA [11] accuracy was 93.00% and ITR found 175.00 bits min^{-1}. While effective, ASS-IISCCA requires a significant amount of labeled data from each subject. Last one is Conv-CA [7] with accuracy was 93.88% and ITR was 190.23 bits min^{-1}. While comparable to the proposed method in performance, Conv-CA is computationally expensive due to its use of canonical correlation analysis (CCA). The proposed method uses transfer learning, convolutional neural networks, and data augmentation to increase efficiency. For a 1s time interval, it obtains 94.69% classification accuracy and an ITR of 193.14 bits min^{-1}.

VII. CONCLUSION

In this research, important issues have been addressed in SSVEP-based BCIs, with an emphasis on using deep learning to increase classification accuracy and information transfer rate (ITR). This method greatly improved performance on a variety of EEG signals using transfer learning to customize a general model to each person. The resilience of the model improved through data augmentation, particularly for shorter signals that are essential in real-time applications. Even with these developments, there are still certain limitations. Although transfer learning and data augmentation work well, they could be improved by investigating more sophisticated deep learning strategies like transformer-based models or attention processes. Furthermore, using advanced augmentation techniques like generative adversarial networks (GANs), could enhance generalization even with sparse data. To increase the applicability of the method, future research should look at cross-dataset and cross-task transfer as well. It will be essential to validate this model in actual BCI systems in order to evaluate its applicability and efficacy. All things considered, this work provides a scalable approach for effective, adaptive BCI systems that successfully handles data limitations and variability.

REFERENCES

[1] W. Yijun, W. Ruiping, G. Xiaorong, and G. Shangkai, "Brain-computer interface based on the high-frequency steady-state visual evoked potential," in *Proceedings. 2005 First International Conference on Neural Interface and Control, 2005.* IEEE, 2005, pp. 37–39.

[2] X. Jiang, B. Cao, L. Ou, Y.-C. Chang, T. Do, and C.-T. Lin, "ifuzzytl: Interpretable fuzzy transfer learning for ssvep bci system," 2024. [Online]. Available: https://arxiv.org/abs/2410.12267

[3] M. N. T. T. Md Rabiul Islam, Md Khademul Islam Molla, "Unsupervised frequency-recognition method of ssveps using a filter bank implementation of binary subband cca," *Journal of Neural Engineering*, 2017.

[4] K. B. L. L. G. S. R. C. D. W. M. T. H. E. C. L. David T. Bundy, Lauren Souders, "Contralesional brain–computer interface control of a powered exoskeleton for motor recovery in chronic stroke survivors," *Stroke*, 2017.

[5] Q. Liu, Y. Jiao, Y. Miao, C. Zuo, X. Wang, A. Cichocki, and J. Jin, "Efficient representations of eeg signals for ssvep frequency recognition based on deep multiset cca," *Neurocomputing*, vol. 378, pp. 36–44, 2020.

[6] S. Mahmood, I. Farhana, M. R. Islam, and M. K. I. Molla, "Frequency recognition of short-time ssvep signal using corrca-based spatio-spectral feature fusion framework," *IEEE Access*, vol. 9.

[7] Y. Li, J. Xiang, and T. Kesavadas, "Convolutional correlation analysis for enhancing the performance of ssvep-based brain-computer interface," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 12, pp. 2681–2690, 2020.

[8] H. Xiong, J. Song, J. Liu, and Y. Han, "Deep transfer learning-based ssvep frequency domain decoding method," *Biomedical Signal Processing and Control*, vol. 89, p. 105931, 2024.

[9] J. Chen, Y. Zhang, Y. Pan, P. Xu, and C. Guan, "A transformer-based deep neural network model for ssvep classification," *Neural Networks*, vol. 164, pp. 521–534, 2023.

[10] W. Ding, A. Liu, L. Guan, and X. Chen, "A novel data augmentation approach using mask encoding for deep learning-based asynchronous ssvep-bci," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 32, pp. 875–886, 2024.

[11] Q. Wei, Y. Zhang, Y. Wang, and X. Gao, "A canonical correlation analysis-based transfer learning framework for enhancing the performance of ssvep-based bcis," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 2809–2821, 2023.