
Advancing SSVEP-Based BCI Speller Using Deep Neural Networks and Letter Frequency Integration

Group 19 - Final Report

Joseph Zhang
Biomedical Engineering
Carnegie Mellon University
Pittsburgh, PA 15213
josephzh@andrew.cmu.edu

Ruiming Zhang
Biomedical Engineering
Carnegie Mellon University
Pittsburgh, PA 15213
ruiming2@andrew.cmu.edu

Kipngeno Koech
College of Engineering
Carnegie Mellon University Africa
BP 6150, Kigali, Rwanda
bkoech@andrew.cmu.edu

David Hill
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213
davidhil@andrew.cmu.edu

Keywords: SSVEP, Brain-Computer Interface, EEG, Deep Learning, Letter Frequency

1 Abstract

Steady-State Visual Evoked Potential (SSVEP) spellers are a promising communication tool for individuals with disabilities. However, they often suffer from low accuracy and poor generalizability to unseen subjects, largely due to the high variability in EEG data. In this study, we propose a hybrid approach combining data augmentation and language modeling to enhance the performance of SSVEP spellers. Using the Benchmark dataset from Tsinghua University, we explore various data augmentation techniques, including frequency masking, time masking, and noise injection, to improve the robustness of deep learning models. Additionally, we integrate a language model (CharRNN) with EEGNet to incorporate linguistic context, significantly enhancing word-level decoding accuracy. Our results demonstrate accuracy improvements of up to 2.9 percent over the baseline, with time masking and language modeling showing the most promise. This work paves the way for more accurate and generalizable SSVEP speller systems, offering improved communication solutions for individuals with disabilities.

2 Introduction

Brain-Computer Interfaces (BCIs) are systems that allow direct communication between the brain and an external device. BCIs have transformative potential, particularly for individuals with severe physical disabilities, offering applications in communication, environmental control, and mobility. Our group is primarily interested in steady-state visually-evoked potentials (SSVEPs), which is a neural signal that occurs when a person looks at flickering stimuli at specific frequencies. This signal is commonly used in a speller task, where a user will try to spell words by looking at a computer screen with a matrix of characters each flickering at specific frequencies, and their SSVEP data is recorded using electroencephalography (EEG). This data is preprocessed and then decoded to determine what the user was looking at, and recently, the use of deep neural networks for decoding SSVEPs has become more viable and effective. However, this process is still slow and prone to errors due to the noisy and highly-variable nature of EEG data. Therefore, our group plans to experiment with various components of the BCI data pipeline to improve the efficacy of SSVEP DNN decoders.

To improve SSVEP BCI performance, our group plans to explore data augmentation and language-based letter frequency modeling. Data augmentation aims to increase the accuracy and stability of a model by creating more data to be used in training which is derived from the original dataset. This new data can be transformed from the dataset using a variety of functions. For our project, we plan to combine the common augmentation of noise addition and Fourier domain random phase shifting to improve our baseline model’s performance.

Another proposed extension is incorporating letter frequency into character classification as the real-world application of these speller tasks would involve writing words and phrases which can be predicted using prior inputs. Beginning with simpler n-grams and progressing to models like Kneser-Ney and RNNs, this language model would dynamically adjust EEGNet’s output, emphasizing linguistically probable characters in noisy conditions.

3 Literature Review

- Guney et al. [2022] introduces a novel DNN architecture for steady-state visual evoked potential (SSVEP)-based brain-computer interfaces (BCIs) that significantly improves information transfer rates (ITRs) for BCI speller systems. The BCI speller task involves looking at various characters/symbols on a grid on a computer screen and attempting to spell words/phrases using just the input of sight, which is very beneficial to people with paralysis. The network involves multiple convolutional layers that captures the spatial and spectral information from the EEG signal. When tested on publicly-available benchmark and BETA datasets, it reached ITRs of 265.23 bits/min on the benchmark dataset and 196.59 bits/min on the BETA dataset, both achieved with just 0.4 seconds of stimulation.
- Liu et al. [2020] introduces the BETA dataset, a large-scale public benchmark for SSVEP-BCIs, and it is used in works such as Guney et al. [2022]. BETA includes EEG recordings from a 40-character speller using a range of flickering frequencies. The dataset contains multiple blocks of trials, with variations in signal quality due to real-world conditions, which makes it more representative of practical applications compared to controlled lab datasets. The paper also compares the performance of several frequency recognition methods (supervised and training-free) on the BETA dataset. Methods like TRCA, CCA, and FBCCA are evaluated to establish performance baselines for the SSVEP classification task.
- Zhang et al. [2024] makes progress in the field of BCI intention decoding for control of unmanned aerial vehicles (UAV’s) and other intelligent vehicles using a lightweight network. There are three main paradigms or signals that have historically been used for this problem: motor imagery (MI), secure typing via BCI system with encrypted feedback (SSVEP), and average evoked potentials (ERP). This work uses both home-grown and public datasets with MI and SSVEP measurements for training, validating, and testing their model. Their network is composed of three main parts: a spatial feature extraction module (a 2D convolution followed by batch normalization), a temporal feature extraction module (a 1D convolution followed by batch normalization), and an attention feature enhancement module (noise adding, pooling, and fully connected layers). Finally, the output is flattened, fed through a few fully connected layers and finally a softmax layer for classification.
- Nie and Wang [2024] gives an overview of the challenges that multiclass classification pose in BCI deep learning. Good performance on datasets with more than 40 categories, such as the Benchmark dataset from Tsinghua University used in the paper, has not yet been achieved. To work towards this goal, the authors propose a novel analysis based on convolutional layers with residual connections. They achieve impressive results which they compare to other popular architectures trained on the same dataset. They achieve 52% accuracy on a 0.2 second window and over 93% accuracy on a 1 second window.
- Ding et al. [2024] introduces an innovative data augmentation technique known as EEG mask encoding (EEG-ME) for deep learning (DL)-based classification in steady-state visual evoked potential (SSVEP)-based brain-computer interfaces (BCIs). This method addresses a common challenge in DL models for BCIs: overfitting caused by the limited availability of electroencephalography (EEG) data. The proposed EEG-ME method improves model generalization by masking portions of the EEG data, thereby encouraging the model to learn more robust features. To validate the effectiveness of EEG-ME, three distinct network architectures were employed: CNN-Former (a hybrid model integrating convolutional

neural networks and Transformer), time-domain-based CNN (tCNN), and a lightweight architecture known as EEGNet. Testing was conducted on publicly available benchmark and BETA datasets. The study demonstrates that EEG-ME substantially improves the average classification accuracy across these architectures and datasets.

- Rostami et al. [2022] explores the enhancement of classification in real-world steady-state visual evoked potential (SSVEP) data for Brain-Computer Interface (BCI) speller systems using deep convolutional neural networks (DCNNs). This study proposes using DCNNs to automatically extract relevant features from noisy SSVEP data to address the issue of noise and artifacts in the signal. The network was pre-trained using a portion of the BETA data and then retrained and tested at the individual subject level. Results indicated a significant improvement in both accuracy and information transfer rate (ITR) after retraining, with a p-value of less than 0.01 for all participants. Specifically, accuracy improved by 25.72 percent, and the ITR increased by 43.10 bits per minute (bpm).
- Wan et al. [2023] presents GDNet-EEG, an attention-aware deep neural network based on group depth-wise convolution for recognizing steady-state visually evoked potential (SSVEP) stimulation frequencies, with potential applications in early glaucoma diagnosis. The model employs group depth-wise convolution to extract temporal and spectral features from EEG signals across brain regions. It also integrates an attention mechanism to highlight essential brain regions for improved feature extraction. The model is validated on two public SSVEP datasets (benchmark and BETA), as well as their combined dataset. With 1-second input signals, GDNet-EEG achieves average classification accuracies of 84.11%, 85.93%, and 93.35% on the benchmark, BETA, and combined datasets, respectively. Compared to baseline models, accuracies improve by 1.96% to 18.2%.
- Lashgari et al. [2020] is a very useful overview paper of various augmentation techniques used in the field of deep learning for EEG. It reviews the common augmentations as well as their effectiveness on different types of problems for BCI-related tasks. It splits the types of data augmentation into six main categories: noise addition, GAN's for generating new data, sliding or overlapping windowing, sampling, Fourier transform transformations, and recombination after segmentation. The authors discuss the current state of the art in each of these techniques and which techniques generally perform better for different tasks such as motor imagery, motor task, and visual task. We plan to use this paper as well as the papers referenced in the article to determine a new augmentation technique.
- Kneser and Ney [1995] introduces an advanced method for smoothing in n-gram language models, known as Kneser-Ney smoothing. This method addresses the issue of estimating probabilities for rare or unseen word sequences in language models by assigning them higher probabilities, which allows the n-gram to be more accurate and generalizable to human speech and writing patterns. This is a very useful system that we can leverage to predict characters to change probabilities in our EEGNet's classification layer.
- Graves [2013] introduced a comprehensive framework for Recurrent Neural Networks (RNNs) in sequence-based tasks, particularly emphasizing their ability to capture temporal dependencies over long sequences. By using a hidden state that evolves as the network processes each time step in the input, RNNs can retain information about previous elements, allowing for context-based predictions of characters, words, and phrases. For our purposes, we want to utilize a character-driven RNN to augment the classification probabilities of specific characters to aid the accuracy of EEGNet.

4 Baseline Model

EEGNet, developed by Lawhern et al. [2018], is a deep learning model specifically designed to analyze electroencephalography (EEG) data for brain-computer interface(BCI). As shown in Figure 1, it is a compact convolutional neural network that uses depthwise and separable convolutions to extract EEG features and perform classification. EEGNet has shown strong performance in various EEG classification tasks, which makes it a reliable baseline for comparison with other models.

Fundamentally, the way EEGNet works is by leveraging convolutions to capture both the spatial and temporal aspects of the EEG data. EEGNet begins with a temporal convolution to capture frequency information. This is achieved by applying F_1 filters of size 1 x 64, where the filter length

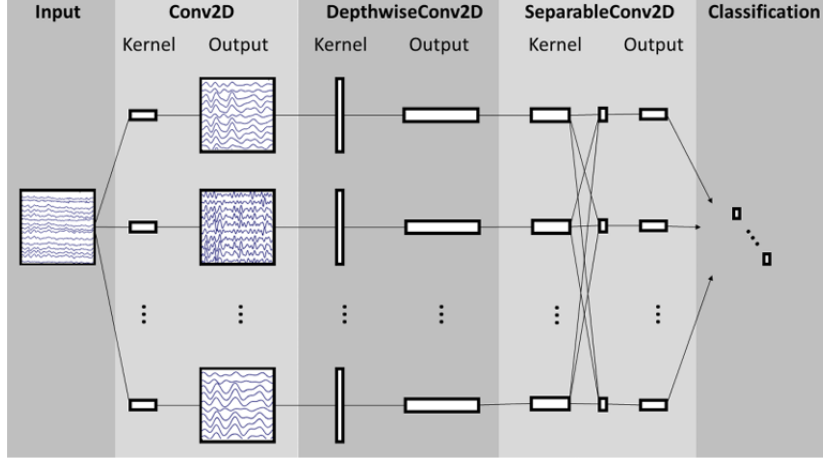


Figure 1: The architecture of EEGNet [Lawhern et al., 2018].

corresponds to half the EEG sampling rate, allowing it to capture frequency information above 2 Hz. The second block uses depthwise convolution to learn spatial filters specific to each temporal filter. For each feature map from the previous layer, EEGNet applies a separate filter across the spatial dimension (channels), enabling EEGNet to learn spatial representations for each frequency band, like for instance, how alpha power is distributed along the cortex. This depthwise convolution has $D \times F_1$ filters, where D is the depth multiplier controlling the number of spatial filters per temporal filter. The next layer uses separable convolution, which combines depthwise convolution (applied independently to each feature map) and pointwise convolution (applied across all feature maps) to reduce parameters and improve feature learning. Note that each layer includes batch normalization, dropout, and activation which is standard for CNNs. The final layer is a classification layer that assigns an array of probabilities for the various classes to the input signal.

In SSVEP classification tasks, EEGNet is widely used as a baseline model to evaluate the performance of new proposed models.[Wan et al., 2023, Ding et al., 2024] For example, Wan et al. [2023] uses EEGNet as one of the baseline models to demonstrate the efficiency of their GNet-EEG in SSVEP classification. The model is validated on two public SSVEP datasets (benchmark and BETA), as well as their combined dataset. With 1-second input signals, Figure 2 shows the classification accuracies over 10-fold cross-validation using a signal length of 0.8 and 1 s. Similarly to these articles, we will be using classification accuracy on the validation set as a way to evaluate the efficacy of our model extensions. Furthermore, evaluating the classification accuracy with SSVEP signals of different time lengths (like 0.8 s versus 1 s), will also be useful as models may perform differently given the size of the input vector. The benefit of using shorter time segments is an increased information transfer rate as the BCI user can look at each character for a shorter amount of time before classification is achieved, which is what we would want from an efficient online BCI system.

Although EEGNet is not the most advanced deep neural network for EEG BCI applications, it is one of the most used and reliable versions. The state-of-the-art models today such as FB-SSVEPformer and GNet-EEG are all based off of EEGNet, and we wanted to use EEGNet as to not confound our proposed model extensions with these other models as they may not be compatible.

In our research, we plan to use EEGNet as a baseline model to evaluate the performance of our architecture.

5 Data

For evaluating the performance of our models, there are several well-established and commonly-used public SSVEP datasets. The Tsinghua SSVEP BCI speller dataset, also known as the Benchmark dataset, includes EEG recordings from 35 participants using a 40-character speller with flickering stimuli ranging from 8 to 15.8 Hz [Wang et al., 2017]. It's widely used to evaluate SSVEP-based BCI systems due to its controlled setup. The BETA dataset, a more recent and larger benchmark,

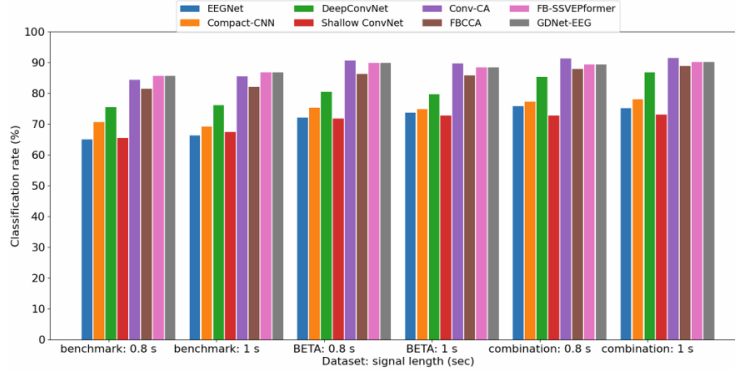


Figure 2: The average classification accuracies obtained by GDNet-EEG and five other baseline models including EEGNet [Wan et al., 2023].

consists of EEG data from 70 participants, also using a 40-character speller, but with additional real-world conditions to better simulate practical applications [Liu et al., 2020]. We will use both of these datasets for our model evaluations.

6 Baseline Implementation

As mentioned in previous sections, we will use EEGNet as our baseline model for SSVEP classification task with benchmark and BETA dataset (See Figure 2).

Model Implementation

The EEGNet model is imported from GitHub (<https://github.com/vlawhern/arl-eegmodels>) [Lawhern et al., 2018]. It is configured with the following parameters: nbclasses=40 specifies 40 output classes, corresponding to the target SSVEP frequencies. The input EEG data has Chans=64 for the 64-channel setup and Samples=250 to represent each trial’s 250 time points. A dropout rate of dropoutRate=0.5 is applied to prevent overfitting. Temporal convolution uses a kernel length of kernLength=64 to capture time-dependent patterns, with F1=8 filters for the first layer. Depthwise convolution is configured with a multiplier D=2 to increase spatial feature extraction, followed by F2=16 filters in the second convolutional layer to refine spatial representations further. The model uses standard dropoutType=’Dropout’ for regularization.

Data Preprocessing

For data preprocessing, EEG signals are filtered using a Chebyshev Type I filter with cutoff frequencies set between 6 and 90 Hz and stopband corners from 4 to 100 Hz as suggested by Wan et al. [2023]. This is done to eliminate biological artifacts like heart beats and eye blinks which have relatively low frequencies and muscle contractions which can have higher frequencies than neural oscillations. Each trial of multi-channel EEG data is then segmented into smaller chunks, with each segment 1 second length as that is how long we want to train the model with (the shorter the better as mentioned previously). The segmentation is also labeled according to the target frequency presented during that trial. Additionally, normalization is applied channel-wise to standardize the EEG data.

Training Hyper-parameters

The model is trained with a batch size of 64 and optimized using the Adam optimizer with a learning rate of 0.001. To robustly assess model performance, we employ ten-fold cross-validation, training each fold for 50 epochs. K-fold cross-validation involves splitting the training data into k-partitions and training a new model for the whole dataset minus one of the partitions, saving that one for validation. This process is then repeated k times, which each partition representing the validation set once. This is useful as BCI speller datasets are often quite small given how difficult it is to collect each trial, so k-fold allows us to maximize the usage of our data. The model’s overall performance is evaluated by calculating the mean accuracy across the 10 folds.

Results

Our baseline model achieved an accuracy of 0.6035 on the benchmark dataset, closely matching the baseline performance reported by Wan et al. [2023] shown in Figure 2. This indicates that our model implementation and data preprocessing are efficient. Due to time constraints, we did not perform additional fine-tuning, which may account for the slightly lower accuracy compared to Wan et al. [2023]. Overall, we have successfully established the pipeline and developed a solid baseline for our research. We are also running the same baseline model on the BETA dataset. However, the model has not finished training so we do not have any numbers to report currently. We aim to achieve a similar closeness to the results reported in Wan et al. [2023].

7 Proposed Model Extensions

7.1 Data Augmentation

One ongoing struggle in the field of BCI is generalizing learned models to individuals not in the training data. Everyone’s brain is wired differently and has various electrical characteristics, all of which make it hard for a model to generalize well. One area of research which is promising on this front is that of creative data augmentation that uses domain knowledge and expertise to multiply the training data to help in generalization. There are many possible augmentation techniques for SSVEP data, but we will list a few here. In Ding et al. [2024], the authors utilize mask encoding to force the model to learn more robust features to noise. In this paper, the authors only mask in the time domain, but it could also be worth looking into masking in the frequency domain. In Mai et al. [2023] a different approach is taken in which the authors build off of previous work in which time-shifted sections are used in training (essentially time-cropping). In this paper, the amount of shifting was determined intelligently for each SSVEP class based on the local similarities of data in a single recording.

In another example, Pan et al. [2024] augment SSVEP data using two methods, both of which are done in the frequency domain after applying an FFT. The first is filter band masking which is essentially applying a notch filter to the data. The second is random phase erasing in which random sections of the frequency data are phase-shifted to a random phase, then the data is returned to the time domain. This is a very interesting concept to us as it attempts to recognize that the time that a signal takes to reach an electrode varies significantly on the person. One weakness we see in this approach is the fact that, in reality, the phase shift of frequency components are not random, but determined by physical characteristics of the brain and skull.

We plan to investigate similar frequency-domain augmentations in our project. Our idea is to come up with a framework in which we can add "frequency-domain noise" to the data in ways that reflect how electric potentials actually propagate through the head. To do this, we will develop a model for how these potentials propagate then vary the parameters of that model and apply a convolution of this filter to the data. We believe this will help the model generalize to unknown subjects as well as improve overall accuracy in the validation dataset.

7.2 Letter Frequency

One area of novel experimentation our group intends to pursue is improving the classification of SSVEPs for the BCI speller task by incorporating letter frequency into the character classification process. Currently, SSVEP-based spellers only use the spectral, temporal, and spatial features of the EEG data to determine a character’s identity, but while that would be the only option if the user is looking at arbitrary stimuli, we know that the speller task contains other useful information. If we assume the user is trying to specifically spell English words, we know that if they first spell the letter 'Q', we can be quite confident that the letter 'U' will immediately follow. Therefore, we can assign a higher weight to the probability of the letter 'U' after confirming a 'Q'. Hypothetically, this may improve the accuracy of the SSVEP speller, especially when a specific trial has low confidence due to noise.

We will test multiple methods for character prediction, increasing in complexity and performance if time allows, which may be unnecessary if a simpler model proves to be the most effective. The simplest model for character prediction is the n-gram, which simply analyzes all n-length character sequences in a given text corpus and then uses conditional probability to predict letters [Shannon, 1948]. One variant of the n-gram is the Kneser-Ney n-gram, which is a language model that can

predict the likelihood of a word or character sequence by adjusting probabilities with a smoothing technique that improves predictions for rare or unseen sequences [Kneser and Ney, 1995]. Unlike basic n-gram models, Kneser-Ney assigns higher probabilities to rare sequences if they occur in varied contexts. There are a variety of other n-gram models that can be looked into but we will begin with this one.

A more advanced model would be the character-level recurrent neural network (RNN), developed by Graves [2013]. This RNN would work by sequentially processing each character in a text, using each character as input while updating its hidden state to capture context from previous characters, allowing the model to predict future letters. At the most advanced level would be large language models using transformers and attention mechanisms, which we will leave as a potential exploration if time permits.

Fundamentally, the way we would incorporate the language module is by using the outputted probabilities to adjust the final classification layer in EEGNet, with probabilities of any character 'C' begin computed using Equation (1), where EEGNet(C) is the base probability from EEGNet and LM(C) is the predicted probability of that character using previous characters. We would test to find the best α to use.

$$P(C) = \alpha * EEGNet(C) + (1 - \alpha) * LM(C) \quad (1)$$

Another issue is the data. While Benchmark and BETA do provide the frequency and phase information of each target so we can determine the true label of each trial, their experimental procedure involved random cued spelling as opposed to spelling words or phrases. However, we can solve this issue by creating our own sentences by concatenating trials. Although this is not the same as the the speller user themselves attempting to spell a sentence, it will still provide a useful benchmark for evaluating the efficacy of language models combined with EEGNet.

8 Work Distribution

By this time, all team members helped design and test the baseline EEG model, as well as conduct research for hypotheses. However, we list below the tasks that each team member led and spent significant time working on.

- Joseph: Imported benchmark data into google drive, tested various implementations of the baseline model, completed literature review and research for letter frequency.
- Ruiming: Implemented data pre-processing for the baseline model and ran the model to get the baseline accuracy for the benchmark dataset.
- David: Completed literature review and research for data augmentation, imported BETA data into google drive, and ran the model to get the baseline accuracy for the BETA dataset (not finished running).
- Kipngeno: Imported the baseline EEGNet model into a shared file and tested various implementation of the baseline model.

9 References

References

- Wenlong Ding, Aiping Liu, Ling Guan, and Xun Chen. A novel data augmentation approach using mask encoding for deep learning-based asynchronous ssvep-bci. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 32:875–886, 2024. doi: 10.1109/TNSRE.2024.3366930.
- Alex Graves. Generating sequences with recurrent neural networks. *arXiv preprint arXiv:1308.0850*, 2013.
- Osman Berke Guney, Muhtasham Oblokulov, and Huseyin Ozkan. A deep neural network for ssvep-based brain-computer interfaces. *IEEE Transactions on Biomedical Engineering*, 69(2): 932–944, 2022. doi: 10.1109/TBME.2021.3110440. URL <https://doi.org/10.1109/TBME.2021.3110440>.

- Reinhard Kneser and Hermann Ney. Improved backing-off for m-gram language modeling. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 181–184. IEEE, 1995.
- Elnaz Lashgari, Dehua Liang, and Uri Maoz. Data augmentation for deep-learning-based electroencephalography. *Journal of Neuroscience Methods*, 346:108885, 2020. ISSN 0165-0270. doi: <https://doi.org/10.1016/j.jneumeth.2020.108885>. URL <https://www.sciencedirect.com/science/article/pii/S0165027020303083>.
- Vernon J Lawhern, Amelia J Solon, Nicholas R Waytowich, Stephen M Gordon, Chou P Hung, and Brent J Lance. Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces. *Journal of Neural Engineering*, 15(5), 2018. doi: 10.1088/1741-2552/aace8c.
- Bixuan Liu, Xiaogang Hu, Yong Wang, Xiangmin Cui, and Xueliang Gao. Beta: A large benchmark database toward ssvep-bci application. *Frontiers in Neuroscience*, 14:630, 2020. doi: 10.3389/fnins.2020.00630. URL <https://www.frontiersin.org/articles/10.3389/fnins.2020.00630/full>.
- Ximing Mai, Jikun Ai, Yuxuan Wei, Xiangyang Zhu, and Jianjun Meng. Phase-locked time-shift data augmentation method for ssvep brain-computer interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31:4096–4105, 2023. doi: 10.1109/TNSRE.2023.3323351.
- Zihao Nie and Raofen Wang. Residual-based convolutional correlation analysis for enhanced models in ssvep classification. In *2024 4th International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, pages 335–339, 2024. doi: 10.1109/ICCECE61317.2024.10504164.
- Yudong Pan, Ning Li, Lianjin Xiong, Yiqian Luo, and Yangsong Zhang. Ssvep data augmentation based on filter band masking and random phase erasing. In Biao Luo, Long Cheng, Zheng-Guang Wu, Hongyi Li, and Chaojie Li, editors, *Neural Information Processing*, pages 483–493, Singapore, 2024. Springer Nature Singapore. ISBN 978-981-99-8138-0.
- Elham Rostami, Farnaz Ghassemi, and Zahra Tabanfar. Improving the classification of real-world ssvep data in brain-computer interface speller systems using deep convolutional neural networks. *Frontiers in Biomedical Technologies*, 9(4):248–254, 2022. doi: 10.18502/ftb.v9i4.10382.
- Claude E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 623–656, 1948.
- Zhijiang Wan, Wangxinjun Cheng, Manyu Li, Renping Zhu, and Wenfeng Duan. Gdnet-eeg: An attention-aware deep neural network based on group depth-wise convolution for ssvep stimulation frequency recognition. *Frontiers in Neuroscience*, 17, 2023. doi: 10.3389/fnins.2023.1160040.
- Yijun Wang, Xiaogang Chen, Xiaorong Gao, and Shang kai Gao. A benchmark dataset for ssvep-based brain-computer interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(10):1746–1752, 2017. doi: 10.1109/TNSRE.2016.2627556. URL <https://doi.org/10.1109/TNSRE.2016.2627556>.
- Wenqi Zhang, Yanjun Qin, and Xiaoming Tao. Decoding brain-controlled intention for uavs and ivs based on lightweight network. *IEEE Internet of Things Journal*, 11(12):21112–21124, 2024. doi: 10.1109/JIOT.2024.3367690.