Currently, there are two main datasets in the field of visual stimulus EEG research: the public visual stimuli EEG dataset from Stanford Digital Repository [1] (utilized in our research) and the ImageNet-EEG Dataset [2]. However, a series of recent studies [3][4] have raised concerns about the experimental paradigm of the ImageNet-EEG dataset, pointing out that the experiment of ImageNet-EEG dataset used a block design so that the same category of stimuli will be continuously presented to the subjects. The stimuli presented in different blocks will be of different categories, leading to an empty high classification accuracy.

In contrast, the public visual stimuli EEG dataset from Stanford Digital Repository used an utterly random order when displaying stimuli. Different categories of stimuli are mixed and randomly presented, avoiding the influence of irrelevant variables caused by block design. Therefore, experimental results obtained from this dataset are more convincible and reliable.

Given the limited availability of visual stimulus EEG datasets, we additionally validated the generalization capability of SAD-VER on the authoritative emotion EEG datasets, SEED [5] and SEED-IV [6].

(1) Dataset description:

The SEED dataset records the EEG signals of 15 subjects while watching movie clips with positive, neutral, and negative emotional content. EEG signals were collected using the 62-channel ESI NeuroScan system at a sampling rate of 1000 Hz, then down-sampled to 200 Hz, and processed using a bandpass filter with a 0-75 Hz range. The SEED-IV dataset is an extension of SEED, which expands the emotional categories to four: happiness, sadness, fear, and neutrality.

(2) Preprocessing workflow:

We extracted the differential entropy (DE) features in five frequency bands ($\delta,\theta,\alpha,\beta,\gamma$) from the EEG signals in the SEED and SEED-IV datasets. Shi et al. [7] had proved that when bandpass filtering is carried out at a 2 Hz step from 2 Hz to 44 Hz, the EEG signals of each sub-band approximately follow the Gaussian distribution, namely, $x \sim N\left(\mu,\sigma^2\right)$. Therefore, the formula for calculating DE is as follows:

$$DE = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \log\left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\right] dx$$

Additionally, following the work presented in Section 3.2 of our original manuscript, we mapped the 62 one-dimensional electrodes to a 9x9 two-dimensional matrix based on their spatial positions in the SEED / SEED-IV datasets, as shown in Figure 2-1. Finally, simple Z-Score normalization was applied to the EEG data.

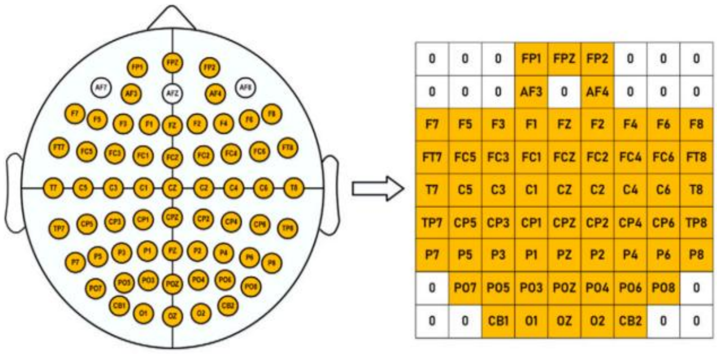

Figure 2-1: Direct Projection Paradigm of EEG in SEED / SEED-IV Datasets

(3) Experimental setup:

Due to differences in the EEG sample data format between the SEED/SEED-IV datasets and the visual stimulus EEG datasets used in the original manuscript, we only adjusted the neural network structures of AV-DPM and STI-Net in SAD-VER. The other modules, hyperparameter settings, and experimental paradigms remained consistent with those in the revised manuscript.

(4) Validation on SEED dataset

We evaluated the performance of SAD-VER on the SEED dataset. Specifically, we compared the classification performance of the STI-Net classification network without the SAD-VER data augmentation framework on the three-class classification task of SEED to demonstrate the performance gain of the proposed method. The results are shown in Figure 2-2. It can be observed that SAD-VER can improve the decoding accuracy by more than 12% on the SEED dataset, with an average improvement of 3.19% in decoding accuracy across the 15 subjects.

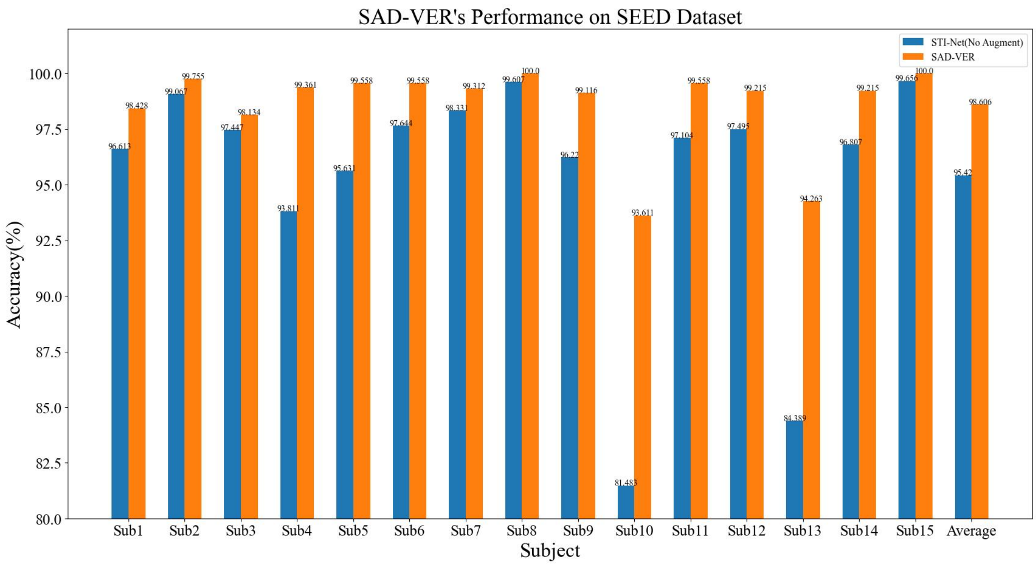

Figure 2-2: SAD-VER's Performance on SEED Dataset

We further assessed the generation quality of SAD-VER when using the SEED dataset, with the results shown in Figure 2-3. A higher FSS is better, and a lower

FSTID is better. We can observe that the FSS values for all subjects exceeded 0.99, demonstrating that SAD-VER can effectively generate high-quality SEED-EEG signals.

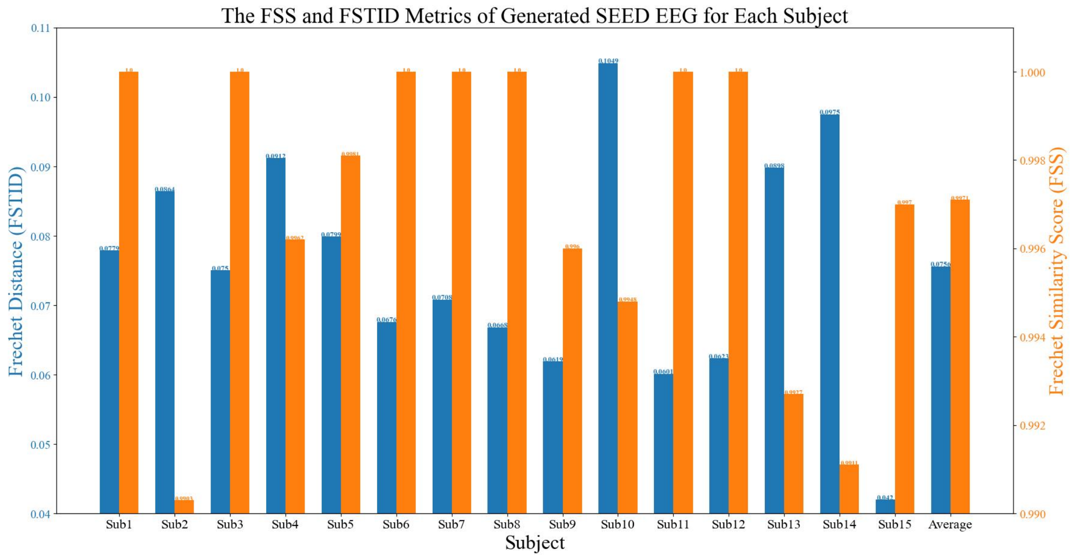

Figure 2-3: The FSS and FSTID metrics for each subject in SEED dataset. Higher FSS values and lower FSTID values indicate better generative quality.

(5) Validation on SEED-IV dataset

We evaluated the performance of SAD-VER on the SEED-IV dataset, and the classification accuracy on the four-class task is shown in Figure 2-4. It can be observed that SAD-VER can improve the decoding accuracy by up to 7.59% on the SEED-IV dataset, with an average improvement of 4% in decoding accuracy across the 15 subjects.

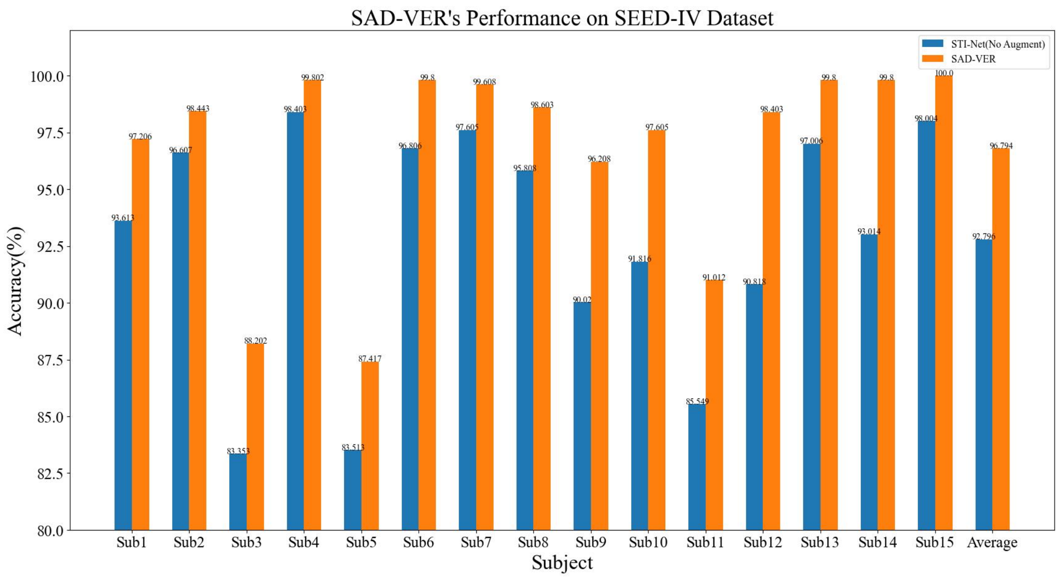

Figure 2-4: SAD-VER's Performance on SEED-IV Dataset

We further assessed the generation quality of SAD-VER when using the SEED-IV dataset, with the results shown in Figure 2-5. We can see that the average

FSS value for the subjects on the SEED-IV dataset is as high as 0.987, indicating that SAD-VER can also generate high-quality EEG signals when using the SEED-IV dataset.

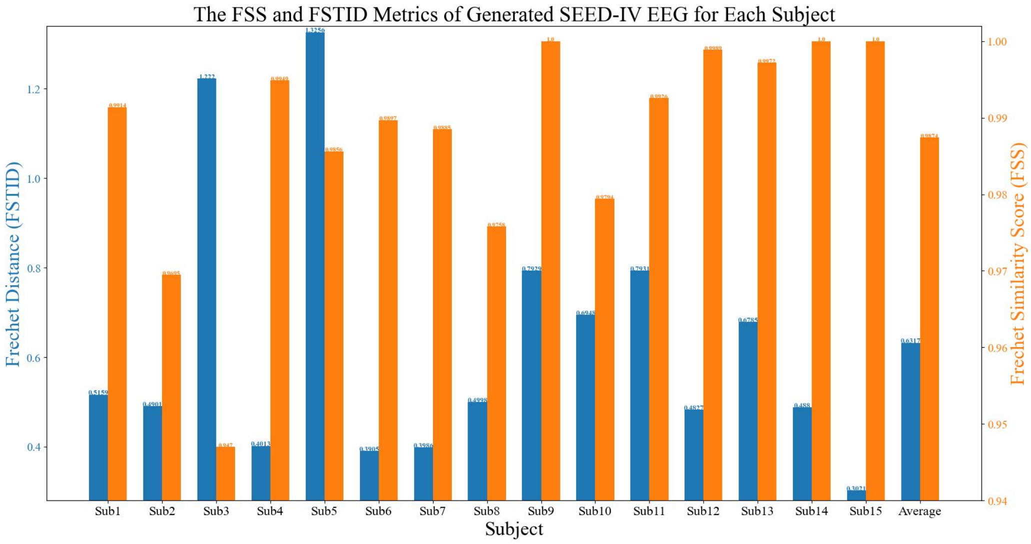

Figure 2-5: The FSS and FSTID metrics for each subject in SEED-IV dataset. Higher FSS values and lower FSTID values indicate better generative quality.

(6) Conclusion

In summary, the experimental results on SEED and SEED-IV demonstrate that SAD-VER can enhance the data through the generation of high-quality emotional EEG signals, significantly improving EEG-based emotion recognition performance. If the classification network STI-Net used in SAD-VER is replaced with a higher-performance classification network, the decoding accuracy should be able to be further improved. These series of experiments highlight the potential of SAD-VER for other types of EEG generation or decoding tasks.

Considering that the focus of this study is primarily on EEG-based visual stimuli research, we have not included the aforementioned experiments in the main text. However, we have made the model files for SAD-VER, when using the SEED and SEED-IV datasets, publicly available on our GitHub page, along with the corresponding instructions. Please refer to the GitHub page here: https://github.com/yellow006/SAD-VER .

Reference:

[1] B. Kaneshiro, M. Perreau Guimaraes, H. S. Kim, A. M. Norcia, P. Suppes, 2015. A representational similarity analysis of the dynamics of object processing using single-trial EEG classification, Plos One, 10(8), e0135697. https://doi.org/10.1371/journal.pone.0135697

[2] C. Spampinato, S. Palazzo, I. Kavasidis, D. Giordano, N. Souly, M. Shah,

2017. Deep Learning Human Mind for Automated Visual Classification, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). https://doi.org/10.48550/arXiv.1609.00344

[3] R. Li, J. S. Johansen, H. Ahmed, T.V. Ilyevsky, R.B. Wilbur, H.M. Bharadwaj, J.M. Siskind, 2020. The Perils and Pitfalls of Block Design for EEG Classification Experiments, IEEE Transactions on Pattern Analysis and Machine Intelligence, 43 (1), 316-333. https://doi.org/10.1109/TPAMI.2020.2973153

[4] H. M. Bharadwaj, R. B. Wilbur, J. M. Siskind, 2023. Still an Ineffective Method With Supertrials / ERPs—Comments on "Decoding Brain Representations by Multimodal Learning of Neural Activity and Visual Features", IEEE Transactions on Pattern Analysis and Machine Intelligence, 45 (11), 14052-14054. https://doi.org/10.1109/TPAMI.2023.3292062

[5] R. Duan, J. Zhu, B. Lu, 2013. Differential entropy feature for EEG-based emotion classification, in: Proceedings of the 6th International IEEE/EMBS Conference on Neural Engineering, 81–84. https://doi.org/10.1109/NER.2013.6695876

[6] W. Zheng, W. Liu, et al., 2019. EmotionMeter: A multimodal framework for recognizing human emotions, IEEE Transactions on Cybernetics, 49 (3), 1110–1122. https://doi.org/10.1109/TCYB.2018.2797176

[7] L. Shi, Y. Jiao, B. Lu, 2013. Differential entropy feature for EEG-based vigilance estimation, in: Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 6627–6630. https://doi.org/10.1109/EMBC.2013.6611075.