

Supplementary Material for Enhancing EEG-Based Schizophrenia Diagnosis with Explainable Multi-Branch Deep Learning

APPENDIX

APPENDIX A: KMUH-SZ-EEG DATASET DESCRIPTION

TABLE A.1: KMUH-SZ-EEG Dataset Overview. The table summarizes the KMUH-SZ-EEG dataset, covering key variables such as subject ID, age of onset, EEG collection age, gender, medication status, and schizophrenia-related symptoms.

SZ Ind.	Onset Age	EEG Exam. Age	Gender	Medication	Symptoms	Hallucination	Delusion	Disorganized	Catatonia	Age/Gender-Matched Control Ind.
S01	38	45	F	1	1	1	1	0	0	C01
S02	59	59	F	0	1	0	1	0	0	C02
S03	20	20	M	1	1	1	0	0	1	C03
S04	33	51	M	1	0	0	0	0	0	C04
S05	31	37	F	1	0	0	0	0	0	C05
S06	19	27	M	0	0	0	0	0	0	C06
S07	46	53	F	1	0	0	0	0	0	C07
S08	26	31	M	1	1	1	0	0	0	C08
S09	23	33	M	1	1	0	1	1	0	C09
S10	41	52	M	1	1	1	0	0	0	C10
S11	49	65	M	1	0	0	0	0	0	C11
S12	21	34	F	1	1	1	0	0	0	C12
S13	15	38	M	1	1	1	1	0	0	C13
S14	30	41	M	1	0	0	0	0	0	C14
S15	18	20	M	1	0	0	0	0	0	C15
S16	21	22	M	1	1	1	0	0	0	C16
S17	32	42	M	1	1	1	0	0	0	C17
S18	31	36	M	1	1	1	1	1	0	C18
S19	26	39	F	1	1	1	0	1	0	C19
S20	36	52	F	1	0	0	0	0	0	C20
S21	21	39	F	1	1	1	0	0	0	C21
S22	28	38	F	1	1	0	1	0	0	C22
S23	18	22	M	1	1	0	1	0	0	C23
S24	42	54	F	1	1	1	0	0	0	C24
S25	18	21	M	1	1	1	1	0	0	C25
S26	17	32	F	1	0	0	0	0	0	C26
S27	33	45	M	1	0	0	0	0	0	C27
S28	31	38	M	1	0	0	0	0	0	C28
S29	29	29	F	1	1	1	1	0	0	C29
S30	20	21	F	1	1	1	1	0	0	C30
S31	32	33	F	1	1	1	1	0	0	C31
S32	37	37	F	1	1	1	1	0	0	C32
S33	20	31	M	1	1	1	1	0	0	C33
S34	29	35	F	0	1	1	1	0	0	C34
S35	31	35	F	1	1	1	1	0	0	C35
S36	21	21	F	1	1	1	1	0	0	C36
S37	29	30	F	1	1	1	1	0	0	C37
S38	29	29	F	1	1	1	1	0	0	C38
S39	27	28	F	1	1	1	1	0	0	C39
S40	20	20	M	1	1	1	1	0	0	C40
S41	34	34	M	1	1	1	1	0	0	C41
S42	38	40	F	1	1	1	0	0	0	C42
S43	25	32	F	1	0	0	0	0	0	C43
S44	17	25	M	1	1	1	1	0	0	C44
S45	45	45	F	1	1	1	1	0	0	C45

SZ Ind.: The subject index assigned to each subject with schizophrenia.

Onset Age: Age of onset based on retrospective review of medical records.

EEG Age: Age at the time of EEG examination.

Gender: Subject's biological sex. (F: Female; M: Male.)

Medication: Whether the subject was on antipsychotic medication at the time of EEG collection. (0: No medication; 1: Medication taken)

Symptoms: Whether the subject exhibited schizophrenia symptoms at initial diagnosis. (0: No; 1: Yes.)

Hallucination: Whether the subject experienced hallucinations. (0: No; 1: Yes.)

Delusion: Whether the subject exhibited delusions. (0: No; 1: Yes.)

Disorganized: Whether the subject showed disorganized behavior. (0: No; 1: Yes.)

Catatonia: Whether the subject demonstrated catatonic behavior. (0: No; 1: Yes.)

Age/Gender-Matched Control Ind.: The subject index of the control subject who is age- and gender-matched with the schizophrenia subject listed in the same row (SZ Ind.).

APPENDIX B: MODEL PARAMETERS, COMPUTATION TIME, AND HYPERPARAMETERS

Table A.2 summarizes the number of trainable parameters, training time, and inference time of the nine models across the IPIN-SZ-EEG and KMUH-SZ-EEG datasets. The number of parameters reflects the total count of trainable weights in each model. Training time refers to the time required to train a single epoch in the first cross-validation fold using a batch size of 16 on an NVIDIA GeForce RTX 4090 GPU. Inference time denotes the duration required to complete forward propagation on the validation set.

Across datasets, the models exhibit a higher parameter count on the KMUH-SZ-EEG dataset due to its bipolar montage input structure. Conversely, training time per epoch was generally longer on IPIN-SZ-EEG, likely due to its greater number of EEG segments. Inference time differences across datasets were negligible. Among the nine models, MBSzEEGNet showed the highest computational load in terms of parameter count, training time, and inference time—attributable to its multi-branch and multi-scale architecture.

While MBSzEEGNet consistently achieved superior classification performance and interpretability, we recognize its increased computational complexity. In real-world clinical applications where latency and hardware resources are limited, this may pose deployment challenges. To address this, future work will explore model compression techniques, including pruning, quantization, and knowledge distillation, to reduce the model’s resource requirements while maintaining performance. Notably, our current analysis indicates that the average inference time remains within sub-second scales, suggesting that with modest optimization, MBSzEEGNet can be adapted for offline diagnostics and potentially real-time or embedded deployment in clinical settings.

TABLE A.2: Model parameters and execution time of the nine DL models in IPIN-SZ-EEG and KMUH-SZ-EEG datasets.

Model	IPIN-SZ-EEG			KMUH-SZ-EEG		
	#Params	Train Time (s/epoch)	Inference (s/segment)	#Params	Train Time (s/epoch)	Inference (s/segment)
Oh-CNN [1]	622	1.09	0.11	637	0.75	0.07
SzHNN [2]	9814	1.37	0.13	9889	0.89	0.08
EEGNet [3]	1858	1.05	0.11	1874	0.69	0.07
SCCNet [4]	6960	0.86	0.09	7242	0.56	0.06
ShallowConvNet [5]	37602	3.39	0.11	39202	2.17	0.07
EEG Conformer [6]	528868	7.61	0.52	530468	4.94	0.33
IFNet [7]	9602	1.10	0.13	9730	0.67	0.08
FBMSNet [8]	13061	3.08	0.25	13349	2.07	0.16
MBSzEEGNet	69430	2.84	0.28	71786	1.85	0.18

Note. #Params: total trainable parameters. Train/inference times measured with batch size 16 on NVIDIA GeForce RTX 4090.

TABLE A.3: MBSzEEGNet’s Optimal hyperparameters (batch size and learning rate) used in 15 repeated runs.

Repeat	IPIN-SZ-EEG		KMUH-SZ-EEG	
	Batch Size	Learning Rate	Batch Size	Learning Rate
1	16	0.0005	16	0.001
2	16	0.001	16	0.001
3	64	0.005	64	0.0001
4	32	0.001	64	0.0005
5	16	0.0005	32	0.0001
6	16	0.0005	64	0.0001
7	64	0.001	64	0.001
8	16	0.0001	16	0.0005
9	64	0.005	16	0.0001
10	32	0.005	64	0.001
11	32	0.005	16	0.0001
12	32	0.005	64	0.0001
13	64	0.005	64	0.0001
14	16	0.0005	16	0.0001
15	16	0.0005	64	0.0005

APPENDIX C: CROSS-DATASET VALIDATION RESULTS

Table A.4 and Table A.5 summarizes the classification performance of the nine DL models under the cross-dataset validation setting. To reduce montage mismatch, we transformed the single-polar recordings of the IPIN-SZ-EEG dataset into bipolar format before cross-dataset evaluation.

The table includes three configurations:

- **Within-IPIN-SZ (Bipolar):** Models trained and tested on the bipolar-transformed IPIN-SZ-EEG dataset using standard 15-fold repetitions.
- **Cross to IPIN:** Models trained on KMUH-SZ-EEG and evaluated on bipolar IPIN-SZ-EEG.
- **Cross to KMUH:** Models trained on the bipolar IPIN-SZ-EEG and evaluated on KMUH-SZ-EEG.

Performance metrics include segment-wise and subject-wise accuracy. Note that due to differing data distributions, electrode montages, and resting-state segment lengths, the overall performance in cross-dataset settings dropped significantly across all models. No statistical significance was observed within the within-IPIN condition (Friedman test $\chi^2 = 9.086$, $p = 0.3351$). These results provide a detailed breakdown of generalizability limitations and supplement the discussion in the main text.

TABLE A.4: Cross-dataset validation performance (%) for within-IPIN-SZ-EEG (bipolar) setting across 9 DL models.

Model	Acc (Seg.)	Acc (Sub.)	Sensitivity	Specificity	Precision	F1-score
Oh-CNN	75.86	79.99	79.99	79.99	80.49	79.90
SzHNN	73.65	79.28	81.90	76.67	79.35	79.80
EEGNet	78.24	81.43	79.99	82.86	82.60	81.21
SCCNet	78.91	81.66	80.95	82.38	82.44	81.50
ShallowConvNet	77.86	81.90	83.81	79.99	81.08	82.25
Conformer	77.04	78.81	82.85	74.76	77.15	79.68
IFNet	78.62	81.43	85.23	77.62	79.46	82.17
FBMSNet	73.32	76.43	77.14	75.71	77.06	76.83
MBSzEEGNet	77.11	79.76	82.86	76.67	78.14	80.32

TABLE A.5: Cross-dataset Validation performance (%) of the nine DL models between IPIN-SZ-EEG and KMUH-SZ-EEG datasets.

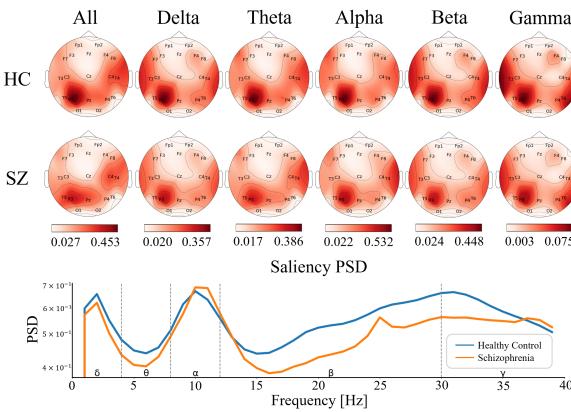
Model	Bipolar-IPIN data evaluated on KMUH model					KMUH data evaluated on Bipolar-IPIN model						
	Acc (Seg.)	Acc (Sub.)	Sensitivity	Specificity	Precision	F1-score	Acc (Seg.)	Acc (Sub.)	Sensitivity	Specificity	Precision	F1-score
Oh-CNN	45.05	43.81	70.95	16.67	45.94	55.73	50.54	50.92	48.99	52.60	47.56	48.15
SzHNN	47.66	45.67	71.52	19.81	47.12	56.72	50.24	50.40	51.43	49.50	47.11	48.75
EEGNet	42.74	39.33	61.62	17.05	42.56	50.31	50.21	51.79	45.64	57.17	48.27	46.89
SCCNet	42.86	39.95	69.33	10.57	43.60	53.49	49.32	51.25	44.95	56.77	47.74	46.22
ShallowConvNet	44.82	42.72	74.09	11.33	45.51	56.37	50.71	51.72	49.03	54.07	48.36	48.63
Conformer	45.70	44.00	76.86	11.14	46.31	57.72	49.89	50.63	51.24	50.10	47.35	49.08
IFNet	41.46	36.05	48.95	23.14	38.87	43.31	47.14	48.39	48.61	48.20	45.09	46.76
FBMSNet	46.53	43.76	79.24	8.28	46.27	58.37	50.25	51.25	53.72	49.10	48.06	50.52
MBSzEEGNet	44.11	41.05	72.67	9.43	44.47	55.15	50.25	51.73	46.25	56.53	48.38	47.02

Note. Acc (Seg.): segment-wise accuracy; Acc (Sub.): subject-wise accuracy.

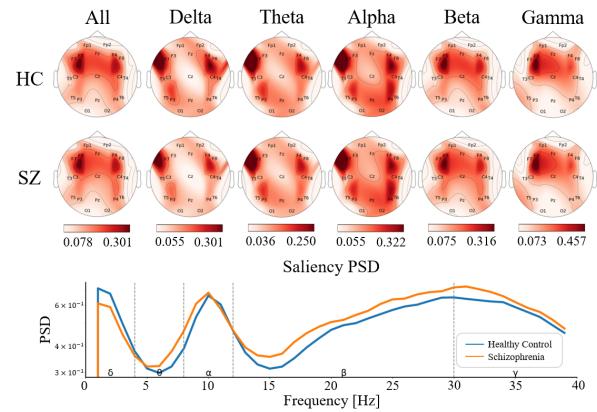
APPENDIX D: ADDITIONAL INTERPRETATION RESULTS

This study used saliency maps to analyze important features for schizophrenia classification based on EEG data, assessing their impact on model outputs. Saliency power spectral density (PSD) reveals activity patterns across different frequency bands in the spectral domain, while saliency PSD-based topomaps uncover associations between frequency bands and brain regions in the spatial domain. Notably, the interpretability results from the IPIN-SZ-EEG and KMUH-SZ-EEG datasets are not directly comparable due to differences in montage formats. Specifically, the bipolar montage used in the KMUH-SZ-EEG dataset, which estimates saliency value based on midpoint channels, introduces variability that may influence the topomap representation.

Fig. A.1-A.5 present the interpretability results of the Oh-CNN, SzHNN, EEGNet, SCCNet, and ShallowConvNet models on the IPIN-SZ-EEG and KMUH-SZ-EEG datasets, highlighting key patterns in feature selection and classification decisions.

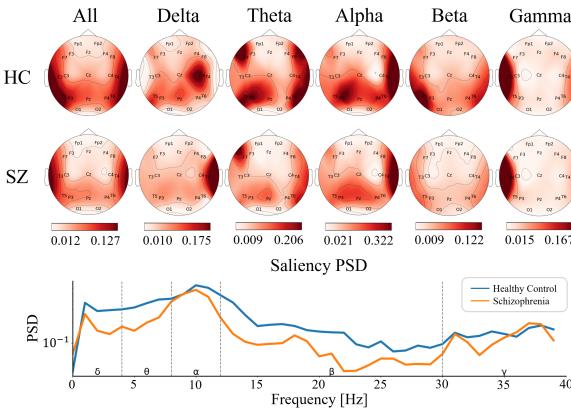


(a) Oh-CNN (IPIN)

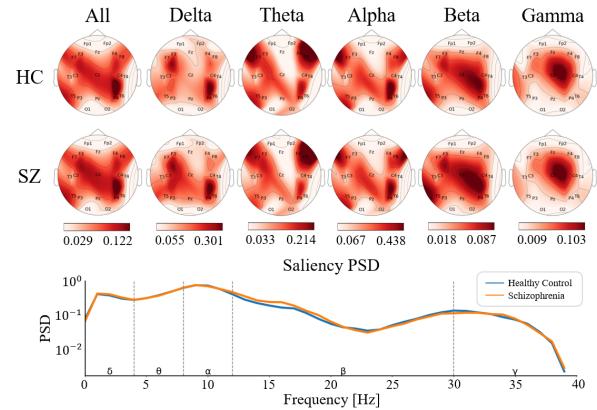


(b) Oh-CNN (KMUH)

Fig. A.1: Saliency PSD and saliency PSD-based topomap of Oh-CNN. (a) Results for the IPIN-SZ-EEG dataset highlight the significance of the delta, alpha, and gamma bands, particularly near the P3 channel. (b) Results for the KMUH-SZ-EEG dataset identify the delta, alpha, and gamma bands, with regions near F7-T3 or F3-C3 crucial for distinguishing healthy controls from schizophrenia patients. HC: healthy control; SZ: schizophrenia patient.

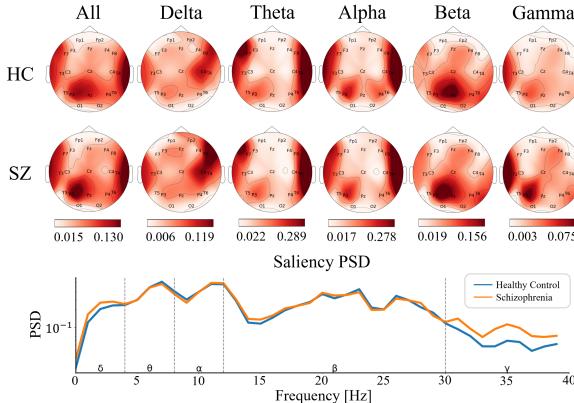


(a) SzHNN (IPIN)

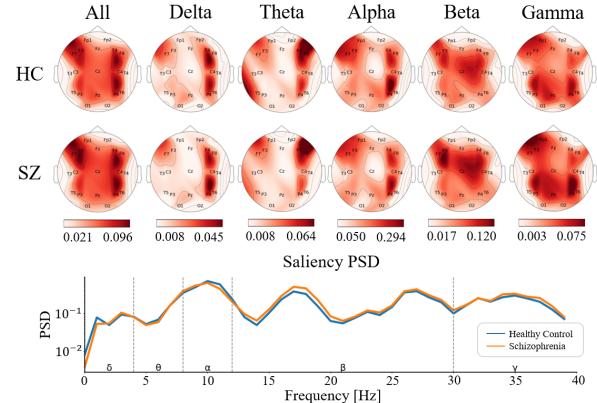


(b) SzHNN (KMUH)

Fig. A.2: Saliency PSD and saliency PSD-based topomap of SzHNN. (a) In IPIN-SZ-EEG dataset, the delta band near the C3, T4 channel and the alpha band near the P3 and T4 channel showed higher importance for classification. (b) In KMUH-SZ-EEG dataset, the C4-P4 region showed higher relevance in delta and alpha bands. HC: healthy control; SZ: schizophrenia patient.

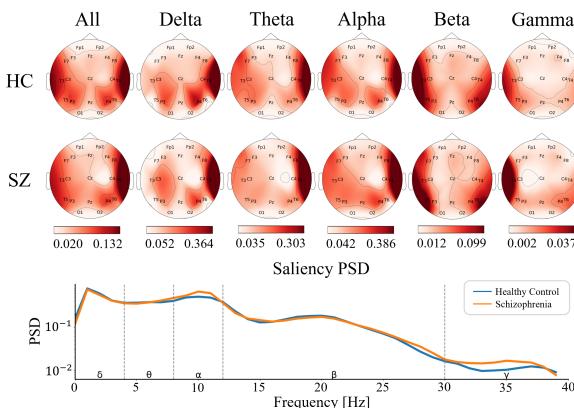


(a) EEGNet (IPIN)

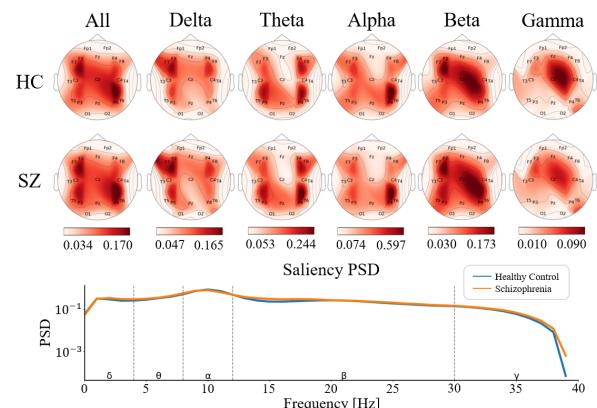


(b) EEGNet (KMUH)

Fig. A.3: Saliency PSD and saliency PSD-based topomap of EEGNet. (a) In the IPIN-SZ-EEG dataset, the result highlights the importance of theta and alpha bands, with contributions from regions near T4. (b) In the KMUH-SZ-EEG dataset, the result shows a relatively greater contribution of the alpha band near the C4-P4 channels to classification. HC: healthy control; SZ: schizophrenia patient.

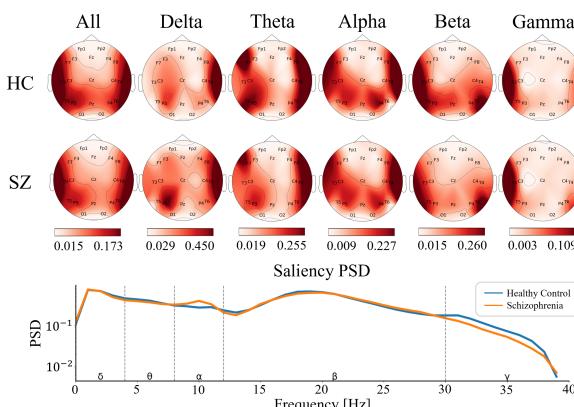


(a) SCCNet (IPIN)

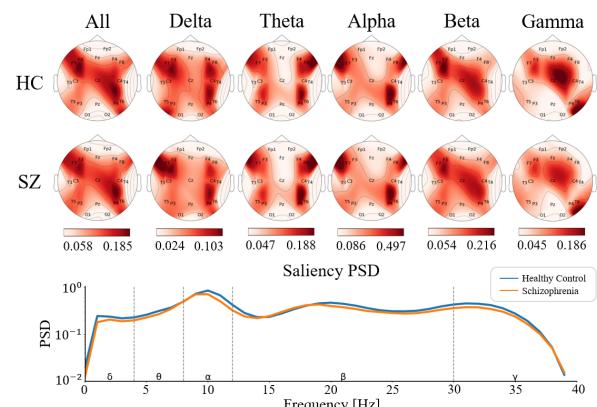


(b) SCCNet (KMUH)

Fig. A.4: Saliency PSD and saliency PSD-based topomap of SCCNet. (a) presents the results for the IPIN-SZ-EEG dataset, where the delta, alpha, and gamma bands were found to be significant, with the T4 and T3 regions playing a key role in classification. (b) shows the KMUH-SZ-EEG results, where the alpha band slightly outperforms other bands, with regions near the C4-P4 channels being more relevant. HC: healthy control; SZ: schizophrenia patient.



(a) ShallowConvNet (IPIN)



(b) ShallowConvNet (KMUH)

Fig. A.5: Saliency PSD and saliency PSD-based topomap of shallowConvNet. (a) For IPIN-SZ-EEG dataset, the delta band at T4 and the beta band between T3, T5 and T6 are significant. (b) For KMUH-SZ-EEG dataset, the result highlights the higher importance of the alpha band in the C4-P4 region for classification. HC: healthy control; SZ: schizophrenia patient.

APPENDIX E: INTERPRETED FEATURE EVALUATION

To validate the causal relevance of saliency-identified features, we conducted a systematic evaluation using the Multi-Domain Adversarial Replacement (mdAR) framework across both spatial and spectral domains [9]. In each domain, we implemented three masking conditions: (1) most relevant features (MoRF), (2) least relevant features (LeRF), and (3) randomly selected features (Rand). Subject-wise classification accuracy was measured under each condition to assess the causal contribution of perturbed features.

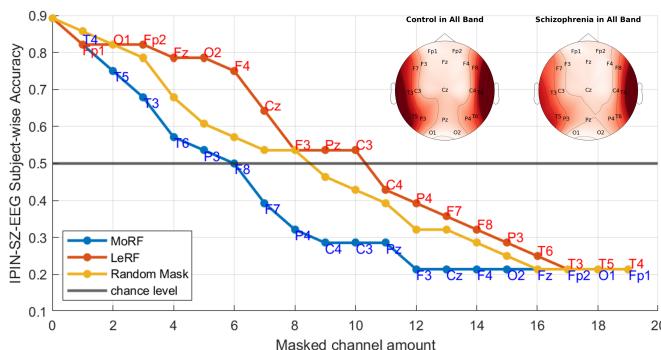
In the spatial domain, we applied Welch's method to compute the power spectral density (PSD) for each EEG channel across trials. These PSDs were averaged across frequency bins to yield a saliency score per channel. The saliency scores were then averaged across subjects to produce a global channel importance ranking. For each condition, a fixed set of channels—either the most salient, least salient, or randomly chosen—was selected for perturbation. We replaced the signal of these selected channels on a trial-by-trial basis using their adversarial counterparts in the time domain. This design preserved temporal and spectral signal structure while selectively altering spatially informative sources.

In the spectral domain, we first computed the PSD for each trial using Welch's method and averaged the values across all subjects to generate a global frequency importance ranking. We then applied FFT to both the original input and its corresponding adversarial version to transform the signals into the frequency domain. For each masking condition, we replaced selected frequency bins—either top-ranked, bottom-ranked, or randomly selected—with their adversarial values. To ensure smooth spectral transitions, we masked each target frequency bin within a ± 0.5 Hz window (e.g., 1 Hz includes 0.5–1.5 Hz). The perturbed frequency-domain signals were then inverse transformed via iFFT back to the time domain for inference.

By comparing model accuracy under different masking conditions, this approach quantitatively validated whether the features highlighted by saliency analysis were indeed causally influential in the model's decision process.

Figure A.6 shows the performance impact of spatial feature masking. In the IPIN-SZ-EEG dataset, replacing the top four spatial channels (T4, T5, T3, T6) caused accuracy to drop significantly from 89.29% to 57.14%, and further masking of P3, F8, and F7 reduced performance below chance level. In the KMUH-SZ-EEG dataset, replacing the top five channels (C4–P4, F7–T3, F4–C4, T3–T5, C3–P3) similarly drove accuracy to chance level.

Figure A.7 displays the effect of spectral masking. In IPIN-SZ-EEG, masking the most salient frequency band (1–3 Hz) resulted in accuracy dropping to chance level. Likewise, in KMUH-SZ-EEG, masking the 9–11 Hz range—identified as most important—caused a similar degradation. Comparisons with least-salient and random masking further confirmed the causal role of high-saliency features.



(a) IPIN-SZ-EEG (spatial)

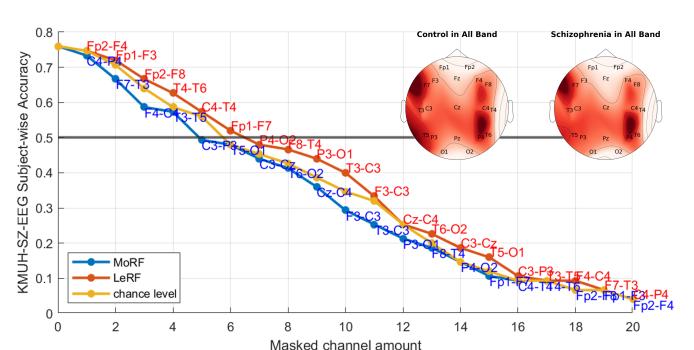


Fig. A.6: Performance degradation after adversarial masking of spatial features. Comparison includes most-relevant, least-relevant, and random masking.

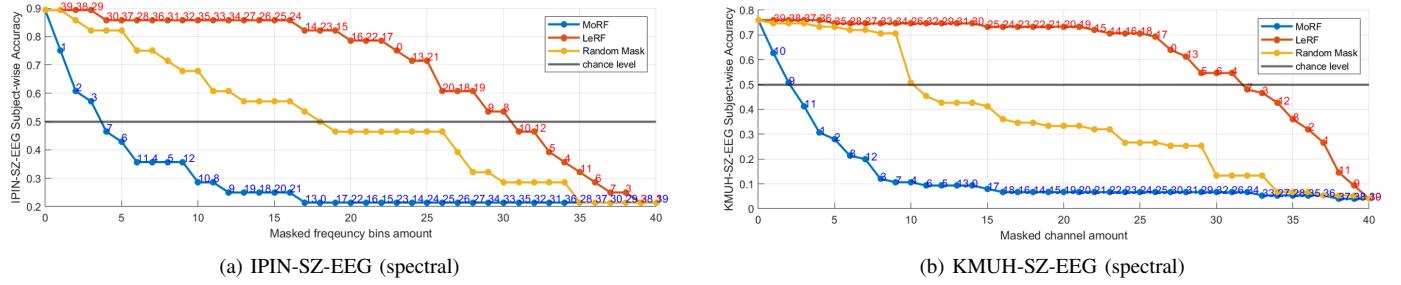


Fig. A.7: Performance degradation after adversarial masking of spectral bands. High-saliency bands cause the most performance drop.

REFERENCES

- [1] S. L. Oh, J. Vicnesh, E. J. Ciaccio, R. Yuvaraj, and U. R. Acharya, “Deep convolutional neural network model for automated diagnosis of schizophrenia using eeg signals,” *Applied Sciences*, vol. 9, no. 14, p. 2870, 2019.
- [2] G. Sharma and A. M. Joshi, “Szhnn: a novel and scalable deep convolution hybrid neural network framework for schizophrenia detection using multichannel eeg,” *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–9, 2022.
- [3] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, “Eegnet: a compact convolutional neural network for eeg-based brain-computer interfaces,” *Journal of neural engineering*, vol. 15, no. 5, p. 056013, 2018.
- [4] C.-S. Wei, T. Koike-Akino, and Y. Wang, “Spatial component-wise convolutional network (scncnet) for motor-imagery eeg classification,” in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2019, pp. 328–331.
- [5] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, “Deep learning with convolutional neural networks for eeg decoding and visualization,” *Human brain mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.
- [6] 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, 2023.
- [7] J. Wang, L. Yao, and Y. Wang, “Ifnet: An interactive frequency convolutional neural network for enhancing motor imagery decoding from eeg,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 1900–1911, 2023.
- [8] K. Liu, M. Yang, Z. Yu, G. Wang, and W. Wu, “Fbmsnet: A filter-bank multi-scale convolutional neural network for eeg-based motor imagery decoding,” *IEEE Transactions on Biomedical Engineering*, vol. 70, no. 2, pp. 436–445, 2022.
- [9] C.-Y. Hsieh and C.-S. Wei, “AIM: Adversarial Information Masking for Evaluating EEG-DL Interpretations,” openreview.net, <https://openreview.net/forum?id=B5i88Tj1nk> (accessed Jun. 11, 2025).