Physiologically Consistent Missing Channel Reconstruction and Cascaded MoE for Impairment Categorization from Visual Electrophysiological Signals

Chenglin Yao[1⋆], Jing Liu[1⋆], Kai Wang[1], Tian Yang[1], Zaidao Han[1], Risa Higashita[2,1✉], and Jiang Liu[1,2✉]

[1] Research Institute of Trustworthy Autonomous Systems & Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China
liuj@sustech.edu.cn
[2] Changchun University, Changchun 130022, China
lisahigashita@gmail.com

Abstract. Visual electrophysiology provides an objective alternative for visual impairment categorization by evaluating stimulus-evoked retinal and neural responses that reflect the functional integrity. However, in practical acquisitions, multichannel recordings frequently suffer from channel missingness caused by sensor failures or environmental interference, leading to degraded robustness in downstream models. While signal reconstruction methods can recover missing channels, conventional approaches often ignore physiological constraints and may introduce hallucinated structures that obscure clinically relevant patterns. In this work, we reformulate it as a physiological consistency-constrained generative task. A domain-specific reconstruction loss is first designed for the GAN-based framework to emphasize stimulus-related waveform regions, preserve multi-scale temporal structure, and constrain magnitude statistics of evoked responses. To address possible failures or noisy cases, an adaptive artifact suppression strategy is further introduced, utilizing a cascaded mixture-of-experts (CMoE) that explicitly models inter-channel dependencies and binocular correlations. Extensive experiments demonstrate that the proposed method substantially improves reconstruction fidelity and provides more accurate impairment gradings.

Keywords: Visual electrophysiology · reconstruction · impairment categorization · physiological consistency · mixture of experts.

1 Introduction

Impairment categorization is an important part of social security systems as it ensures fair and accurate support based on individual needs. While current visual

⋆ Equal Contribution.

impairment categorization standards [3] relying on acuity and visual field measurements are heavily dependent on subjective responses, they are vulnerable to physiological and psychological interference, which results in inaccurate results and even fraudulent practices, such as exaggerating impairment to gain improper benefits. Conversely, visual electrophysiology offers a more objective and quantifiable alternative for visual impairment categorization. With the progression of visual impairments, neural dysfunction triggers consistent and measurable changes in such signals, including delayed evoked potentials and attenuated action potential amplitudes [2, 19]. By capturing and quantifying these changes, visual impairment can be evidenced in the structural characteristics of visual electrophysiological signals, enhancing its diagnostic reliability.

Visual electrophysiology for the recognition of visual dysfunction is an emerging research field. Recently, it has been extensively applied in diagnosing various vision-related diseases [4, 5] and neurodevelopmental disorders [9]. Some preliminary studies [1, 12, 27] have attempted to use this technology to evaluate overall visual function in patients with various pathologies, classifying them into healthy and impaired groups. Despite its promising prospects, an important challenge remains. In practical signal acquisition, multichannel recordings are frequently affected by channel missingness due to sensor malfunctions, motion artifacts, or subject-dependent constraints. Such missingness inevitably leads to the loss of discriminative features and induces a mismatch in input distributions across samples. As a consequence, learning-based models are often sensitive to missing-channel patterns, exhibiting degraded performance.

To mitigate the adverse effects of missing modalities, existing studies have mainly followed two representative paradigms. The first focuses on learning missing-robust encoders or decoders, e.g., through mixture-of-experts architectures [15] or missing-invariant representations [17, 28, 30], where the model is explicitly encouraged to reduce its dependency on any single modality. The second paradigm aims to recover the missing signals explicitly via reconstruction [8], transforming the original task into one with complete inputs. Owing to its ability to exploit population-level priors learned from training data, signal imputation has been widely adopted in practice, with methods ranging from naive [26] and statistical approaches [16] to machine learning-based [18, 20, 29] and generative models, including GANs [13, 14] and diffusion models [7, 25]. Despite their effectiveness, missing channel completion inherently introduces information beyond the observable sample, which is not guaranteed to be physiologically consistent. Such synthesized signals may suffer from hallucinated structures, potentially obscuring clinically critical patterns.

Motivated by the observation that visual electrophysiological signal reconstruction is not merely a data completion problem but a physiological consistency-constrained generation task, we propose a physiologically consistent reconstruction paradigm that explicitly embeds electrophysiological priors into the learning objective. Our approach is grounded in the biophysical characteristics of visual evoked responses, where a domain-specific reconstruction loss is proposed that adaptively emphasizes the fidelity of stimulus-related waveform regions through

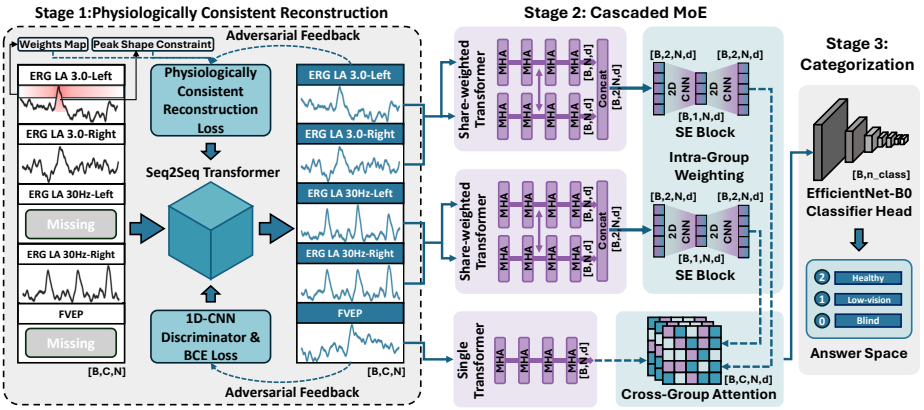

Fig. 1. Overview of the proposed visual impairment categorization framework. MHA referes to multi-head attention.

peak-aware weighting and amplitude characteristics matching. In addition, to suppress harmful hallucinations introduced during signal synthesis, we propose an adaptive filtering strategy operating at the classification head. A cascaded mixture-of-experts (CMoE) mechanism is designed to dynamically attenuate channels exhibiting unreliable or suspicious structures. The primary level gating is performed mode-wise to dynamically prioritize modes carrying high-fidelity semantic information. Then, signals acquired from the left and right eyes often form Siamese channel pairs that share environmental interference and failure patterns, leading to highly synchronized degradation. The cascaded MoE explicitly models this inter-channel dependency at the secondary level gating, enabling coordinated gating across correlated channels and preventing the propagation of parallel noise or hallucinated structures.

In summary, our contributions are: 1) a signal reconstruction paradigm that treats missing channel recovery as a physiology-constrained generation task and enforces waveform, amplitude, and evoked-response fidelity through a domain-specific reconstruction loss; 2) an adaptive artifact suppression strategy using a cascaded mixture-of-experts, explicitly modeling inter-channel and Siamese-channel dependencies to dynamically attenuate unreliable channels; and 3) extensive experiments that demonstrate significant improvements in the proposed method for visual impairment categorization compared to conventional methods.

2 Methodology

2.1 Overview of the Proposed Method

We address the channel missingness by formulating a three-stage framework visualized in Fig. 1, where missing channel reconstruction serves as a strategy pathway to support the primary task of visual impairment categorization. The input

consists of observed multichannel electrophysiological sequences with a missing channel mask, and the outputs include both reconstructed full signals and categorical predictions. In the first stage, a GAN-based module is trained to synthesize missing channels: a sequence-to-sequence Transformer as the generator and a 1D CNN as the discriminator. This stage uses complete electrophysiological recordings without visual impairment labels and is optimized using the proposed physiologically consistent reconstruction loss. The reconstruction objective explicitly incorporates electrophysiological priors, emphasizing stimulus-related waveform regions. In the second stage, the reconstructed signals are used for visual impairment categorization. To suppress hallucinated or unreliable structures introduced during reconstruction, an adaptive filtering mechanism is integrated via a cascaded mixture-of-experts (CMoE) architecture. Siamese channel pairs from the left and right eyes are first processed by share-weighted Transformers with element-wise weighting, followed by inter-channel weighted feature fusion through cross-attention and convolution. The fused representations are finally passed to an EfficientNet-based classification head [24] to produce impairment predictions.

2.2 Physiologically Consistent Signal Reconstruction

Visual electrophysiology recordings are frequently affected by partial channel missingness in real-world acquisitions. As a result, downstream analysis models are exposed to incomplete observations, which can significantly degrade performance if not well addressed. Most existing signal reconstruction methods are designed for general time-series data. These approaches [8, 21, 25] typically do not explicitly incorporate domain-specific priors. From a physiological perspective, electrophysiological responses are governed by the dynamics of resting and action potentials. During most of the recording interval, the signal remains near a stable trough corresponding to the resting membrane potential. Upon visual stimulation, a short-duration action potential is elicited, producing a pronounced peak or peak-valley complex. Importantly, signal fluctuations in the temporal neighborhood surrounding the stimulus response cannot be interfered with, reflecting the latent time needed for depolarization. These properties imply that both waveform shape and amplitude statistics around the stimulus responses are critical for high-quality signal reconstruction.

Based on these observations, we propose a physiologically consistent reconstruction loss for domain-specific missing channel synthesis. By explicitly reweighting reconstruction errors near physiologically meaningful regions and constraining the local amplitude distribution, the model is encouraged to generate signals that are not only visually similar but also statistically consistent with real electrophysiological responses. Formally, let $\boldsymbol{y} \in \mathbb{R}^N$ denote the real signal and $\hat{\boldsymbol{y}} \in \mathbb{R}^N$ the reconstructed signal. We first estimate a baseline mean μ and standard deviation σ from trough-dominated regions of $\boldsymbol{y}$. A soft physiological weight map $\boldsymbol{w}(t)$ is then defined as:

$$\boldsymbol{w}(t) = 1 + \alpha \cdot \sigma\big(\beta(|\boldsymbol{z}(t)| - \tau)\big), \quad \boldsymbol{z}(t) = \frac{\boldsymbol{y}(t) - \mu}{\sigma}, \tag{1}$$

where $\sigma(\cdot)$ denotes the sigmoid function. This weighting amplifies reconstruction penalties near stimulus responses while leaving resting segments weakly constrained. A weighted multi-scale reconstruction loss is computed across several temporal resolutions:

$$\mathcal{L}_{\text{rec}} = \sum_{s} \left(\lambda_1 \|\boldsymbol{w}_s \odot (\hat{\boldsymbol{y}}_s - \boldsymbol{y}_s)\|_1 + \lambda_2 \|\boldsymbol{w}_s \odot (\hat{\boldsymbol{y}}_s - \boldsymbol{y}_s)\|_2^2 \right), \tag{2}$$

where s indexes different downsampling scales. To further encode physiological priors, we model the signal distribution around the stimulus response as a Gaussian distribution. A differentiable soft localization is used to identify the response center, and a temporal window is constructed around it. Within this window, we match the statistics between real and reconstructed signals:

$$\mathcal{L}_{\text{gauss}} = |\mu_{\hat{\boldsymbol{y}}} - \mu_{\boldsymbol{y}}| + |\sigma_{\hat{\boldsymbol{y}}} - \sigma_{\boldsymbol{y}}|. \tag{3}$$

In addition, a soft response amplitude constraint enforces the consistency of the evoked response magnitude:

$$\mathcal{L}_{\text{amp}} = |A_{\hat{\boldsymbol{y}}} - A_{\boldsymbol{y}}|. \tag{4}$$

The final reconstruction objective is a weighted combination of these terms:

$$\mathcal{L}_{\text{physio}} = \mathcal{L}_{\text{rec}} + \lambda_g \mathcal{L}_{\text{gauss}} + \lambda_p \mathcal{L}_{\text{amp}}. \tag{5}$$

By jointly modeling waveform shape, amplitude, and local statistical structure in a physiologically informed manner, the proposed reconstruction strategy produces electrophysiological signals that better align with underlying neural response mechanisms.

2.3 Cascaded Mixture-of-Experts for Signal Recognition

Although signal completion can improve model robustness in the presence of missing channels, it often introduces hallucinated structures that can mislead downstream classifiers. Furthermore, even when signals are not missing, they may suffer from noise or partial corruption, making some channels less reliable. To address this, we propose to adopt a Mixture-of-Experts (MoE) mechanism that adaptively modulates the contribution of each channel. However, traditional MoE designs [15] typically treat each channel as an independent expert. Such methods neglect the known structural correlations between Siamese channels, such as left and right eye recordings collected simultaneously under the same stimulus. In these cases, signal quality degradation in one eye often co-occurs with degradation in the other. Ignoring this relationship limits the MoE's ability to make globally consistent decisions.

To explicitly model such intra-pair dependencies, we introduce a Cascaded MoE (CMoE) framework that performs gating in two hierarchical stages: 1) Intra-Group Gating with Transformer and Squeeze-and-Excitation (SE) Module. To model the varying reliability across electrophysiological channels, we first

group Siamese signals and encode them using a shared time-series Transformer to extract temporal features. The encoded features are element-wise weighted through an SE [6] block. This design enables the network to enhance consistent patterns. 2) Cross-Group Attention for Multi-Modal Fusion. The gated group-level representations are further refined through a cross-group attention mechanism that models inter-modality dependencies. This stage allows the network to capture correlations and contextual reinforcement across different signal types. Multi-head attention is performed across the group axis, followed by residual normalization and linear projection to ensure stable and expressive fusion. The final representation is passed to an EfficientNet head, which projects the features into the answer space.

3 Experiments

3.1 Experimental Settings

We constructed a multimode visual electrophysiology dataset, SPECCU, for generalized visual impairment assessment, collected at a special education college in China. Participants were primarily undergraduates, covering a wide spectrum of ocular and optic nerve disorders, including glaucoma, cataracts, retinitis pigmentosa, macular degeneration, etc. Following [27], subjects were categorized into three clinically and socially relevant classes, i.e., blind, low-vision, and healthy, where the blind class corresponds to the Chinese visual disability certificate level 1, and the low-vision class corresponds to levels 2-4. Visual electrophysiology acquisition comprised ERG and VEP sessions using the HE-2000 system (Tomey, Nagoya, Japan). ERG followed the ISCEV guideline [22] and included two light-adapted protocols: LA 3.0 and LA 30Hz recorded bilaterally. VEP employed a full-field flash stimulus following the ISCEV guideline [23], producing one sequence per subject. There are 2,294 samples in total, in which 805 samples have completed channels, while the remaining suffer channel missingness.

All experiments were conducted using the standard 5-fold cross-validation. For the physiologically consistent reconstruction loss, we set $\lambda_1 = 1.0$, $\lambda_2 = 0.5$, $\alpha = 4$, $\beta = 3$, $\lambda_g = 0.2$, and $\lambda_p = 0.5$. The sequence-to-sequence Transformer imputer consisted of 4 encoder layers with an embedding dimension of 192. For downstream visual impairment categorization, an EfficientNet-B0 head was adopted. Models were trained for up to 50 epochs using an initial learning rate of 1e-4. All experiments were implemented on a workstation equipped with a single NVIDIA RTX 2080Ti GPU. The reported metrics include accuracy (ACC), recall, F1 score, and Quadratic Weighted Kappa (QWK). Recall was chosen as the primary metric. We additionally report the median balanced accuracy (BACC) across the five folds to maintain comparability with prior works [11, 27].

3.2 Comparison on the SPECCU dataset

We first compare representative methods that have been designed for or applied to visual electrophysiological signals on the private SPECCU dataset, including

Table 1. Comparison of existing methods on the private SPECCU dataset. The best-performing method is in **bold** and the second best is underlined.

Methods	ACC	BACC	Recall	F1 Score	QWK
VMamba [27]	0.8365±0.0168	0.6662	0.6675±0.0220	0.6687±0.0238	0.7353±0.0244
GiapDNN [4]	0.8396±0.0200	0.6859	0.6814±0.0160	0.6715±0.0159	0.7814±0.0215
gMLP [9]	0.8579±0.0071	0.7044	0.6995±0.0138	0.7007±0.0145	0.8018±0.0125
TST [10]	0.8588±0.0087	0.7098	0.7022±0.0147	0.6898±0.0220	0.8190±0.0274
MICE [20]	0.8557±0.0095	0.7067	0.7029±0.0088	0.6950±0.0178	0.7959±0.0243
SimMLM [15]	0.8727±0.0106	0.6988	0.7037±0.0250	0.7234±0.0228	0.7919±0.0234
Masking [28]	0.8570±0.0134	0.7198	0.7184±0.0269	0.7026±0.0238	0.8168±0.0233
CGAN [14]	0.8627±0.0121	0.7203	0.7222±0.0164	0.7091±0.0285	0.8096±0.0331
TabTransformer [8]	0.8605±0.0140	0.7193	0.7167±0.0230	0.6978±0.0497	0.8179±0.0241
CSDI [25]	0.8614±0.0072	0.6994	0.7052±0.0218	0.7065±0.0246	0.8102±0.0226
Proposed	**0.8731±0.0203**	**0.7432**	**0.7450±0.0331**	**0.7408±0.0426**	**0.8273±0.0136**

the CNN-based method GiapDNN [4], the transformer-based method time-series transformer (TST) [10], and sequence-aware models (gMLP [9], VMamba [27]). These methods were reproduced to adapt to the target task. They either dropped samples with missing channels or performed naive padding to missing channels. As the blue region of Table 1 shows, the proposed method achieves the best performance across all metrics, with clear gains in ACC, BACC, recall, F1 score, and QWK, indicating improved categorization power under realistic acquisition conditions.

To better demonstrate our contributions to addressing missing channels, we then compare representative approaches for handling missing signals, including statistical imputation (MICE [20]), missingness-invariant representation learning (SimMLM [15], Masking [28]), and generative modeling (CGAN [14], TabTransformer [8], CSDI [25]). These methods were taken from their original repositories. During the reconstruction phase, the self-supervision learning strategy was adopted, with the grading labels invisible to the models. As the green region of Table 1 shows, the proposed method still achieves the best performance across all metrics demonstrating superior robustness to missing channel patterns and more reliable visual impairment categorization. It is interesting to observe that the complex signal reconstruction method does not guaranty a consistent performance gain. For example, CSDI fails to outperform TST in terms of BACC, as the generated signals from the diffusion-based model may contain class-irrelevant or spurious information, introducing hallucinations that mislead the model.

To gain more in-depth insight, we visualize the reconstructed FVEP signal in Fig. 2 as a representative example. The results reveal that MICE and CSDI fail to recover the prominent P2 wave at around 100 ms, while GAN does not adequately reconstruct the N2 wave at around 80 ms. Both are critical components reflecting visual functionality. In contrast, our proposed method successfully preserves both the N2 and P2 components with high fidelity, demonstrating its superior physiological consistency.

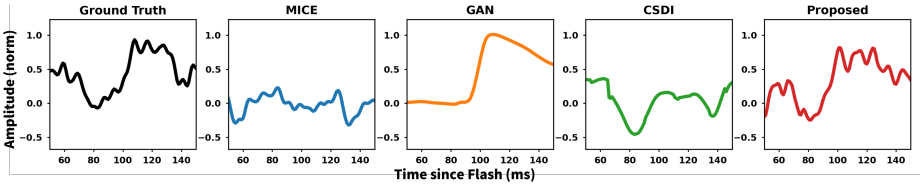

Fig. 2. Reconstruction performance of generative methods with the ground truth.

Table 2. Comparison of existing methods on the public OculusGraphy dataset.

Methods	ACC	BACC	Recall	F1 Score	QWK
VMamba [27]	0.8856±0.0158	0.8761	0.8820±0.0168	0.8767±0.0165	0.7536±0.0328
GiapDNN [4]	0.9125±0.0257	0.9115	0.9123±0.0245	0.9063±0.0253	0.8129±0.0502
gMLP [9]	0.8882±0.0403	0.8782	0.8720±0.0355	0.8773±0.0406	0.7547±0.0814
TST [10]	0.9173±0.0157	0.9069	0.9134±0.0184	0.9103±0.0146	0.8212±0.0288
MICE [20]	0.9148±0.0135	0.9226	0.9194±0.0129	0.9080±0.0120	0.8164±0.0238
SimMLM [15]	0.9075±0.0294	0.9115	0.9139±0.0327	0.9018±0.0302	0.8041±0.0602
Masking [28]	0.9125±0.0277	0.9186	0.9122±0.0263	0.9062±0.0280	0.8125±0.0559
CGAN [14]	0.9028±0.0338	0.9186	0.9105±0.0263	0.8975±0.0325	0.7963±0.0635
TabTransformer [8]	0.9124±0.0158	0.9183	0.9126±0.0168	0.9057±0.0165	0.8118±0.0328
CSDI [25]	0.9076±0.0258	0.9186	0.9123±0.0196	0.9020±0.0244	0.8047±0.0481
Proposed	**0.9319±0.0195**	**0.9388**	**0.9300±0.0222**	**0.9263±0.0202**	**0.8528±0.0404**

3.3 Comparison on the OculusGraphy Dataset

We evaluated the multicenter robustness of the proposed method on the public OculusGraphy dataset, allocating three ERG response modes, i.e., scotopic, maximum, and photopic. Samples exhibit diverse missing patterns, in which 109 of 411 have channel missingness. The results are shown in Table 2. Our method outperforms all competing approaches in multi-center validation, demonstrating its robustness in recognizing visual physiological signals under conditions of missingness.

3.4 Ablation Study

An ablation study was performed to validate the effectiveness of both proposed components, i.e., Physiologically Consistent Loss and Cascaded Mixture-of-Experts, by evaluating performance across all test data as well as stratified subsets with complete and missing signal channels. As Table 3 illustrates, the results indicate that each component individually improves performance, and their combination yields the most significant gains. Notably, the improvement is significantly larger in the missing-channel subset compared to the complete-channel subset, demonstrating that our method effectively mitigates the adverse impact of channel missingness.

Table 3. Ablation study of the proposed components across different data subsets (All, Complete, Incomplete) using BACC, recall, and F1 score.

Components		BACC			Recall			F1 Score		
Physio Loss	CMoE	All	Complete	Missing	All	Complete	Missing	All	Complete	Missing
✗	✗	0.7193	0.7517	0.6623	0.7167	0.7410	0.6760	0.6978	0.7164	0.6581
✓	✗	0.7334	0.7374	0.7048	0.7338	0.7431	0.7058	0.7207	0.7301	0.6931
✗	✓	0.7250	0.7477	0.6654	0.7245	0.7449	0.6770	0.7186	0.7361	0.6716
✓	✓	**0.7432**	**0.7673**	**0.7166**	**0.7450**	**0.7528**	**0.7291**	**0.7408**	**0.7464**	**0.7236**

4 Conclusion

This work addresses missing-channel challenges in visual electrophysiology when grading impairments via the proposed physiologically consistent signal reconstruction paradigm. By embedding biophysical priors into a domain-specific reconstruction loss and integrating an adaptive artifact suppression mechanism with cascaded mixture-of-experts (CMoE) decoding, the proposed framework produces reconstructions that are both numerically accurate and physiologically plausible. Extensive experiments demonstrate improved generalization ability and categorization performance under realistic acquisition conditions.

Acknowledgments. This work was supported in part by National Key R&D Program of China under Grant 2024YFE0198100 and 2024YFC2510800.

Disclosure of Interests. The authors have no competing interests to declare.

References

1. Albasu, F., Kulyabin, M., Zhdanov, A., Dolganov, A., Ronkin, M., et al.: Electroretinogram analysis using a short-time Fourier transform and machine learning techniques. Bioengineering **11(9)**(866) (2024)
2. Barton, J.L., Garber, J.Y., Klistorner, A., Barnett, M.H.: The electrophysiological assessment of visual function in multiple sclerosis. Clinical Neurophysiology Practice **4**, 90–96 (2019)
3. Dandona, L., Dandona, R.: Revision of visual impairment definitions in the international statistical classification of diseases. BMC Medicine **4(1)**(7) (2006)
4. Giap, B.D., Likosky, K., Srinivasan, K., Khan, N.W., Nallasamy, N.: Automated abnormality detection in patient retinal function: A deep learning-powered electroretinogram analysis system. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). pp. 1–4 (2024)
5. Hassankarimi, H., Noori, S.M.R., Jafarzadehpour, E., Yazdani, S., Radinmehr, F.: Analysis of pattern electroretinogram signals of early primary open-angle glaucoma in discrete wavelet transform coefficients domain. International Ophthalmology **39**, 2373–2383 (2019)
6. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 7132–7141 (2018)

7. Jiang, Z., Dai, W., Wei, Q., Qin, Z., Li, K., Zhang, L.: EEG-DIF: Early warning of epileptic seizures through generative diffusion model-based multi-channel EEG signals forecasting. In: ACM International Conference on Bioinformatics, Computational Biology and Health Informatics. No. 71 (2024)
8. Khan, M.A.: A comparative study on imputation techniques: Introducing a transformer model for robust and efficient handling of missing EEG amplitude data. Bioengineering **11(8)**(740) (2024)
9. Kulyabin, M., Constable, P.A., Zhdanov, A., Lee, I.O., Thompson, D.A., Maier, A.: Attention to the electroretinogram: Gated multilayer perceptron for ASD classification. IEEE Access **12**, 52352–52362 (2024)
10. Kulyabin, M., Kremers, J., Holbach, V., Maier, A., Huchzermeyer, C.: Artificial intelligence for detection of retinal toxicity in chloroquine and hydroxychloroquine therapy using multifocal electroretinogram waveforms. Scientific Reports **14(1)**(24853) (2024)
11. Kulyabin, M., Zhdanov, A., Dolganov, A., Maier, A.: Optimal combination of mother wavelet and AI model for precise classification of pediatric electroretinogram signals. Sensors (5813) (2023)
12. Kulyabin, M., Zhdanov, A., Dolganov, A., Ronkin, M., Borisov, V., Maier, A.: Enhancing electroretinogram classification with multi-wavelet analysis and visual transformer. Sensors **23(21)**(8727) (2023)
13. Lee, W., Lee, J., Kim, Y.: Contextual imputation with missing sequence of EEG signals using generative adversarial networks. IEEE Access **9**, 151753–151765 (2021)
14. Leng, Q., Liu, W., Meng, X.: Spectral clustering augmentation pseudo-label conditional generation adversarial imputation networks for multivariate time-series missing data. In: International Conference on Systems and Informatics (ICSAI). pp. 1–6 (2023)
15. Li, S., Chen, C., Han, J.: SimMLM: A simple framework for multi-modal learning with missing modality. In: IEEE/CVF International Conference on Computer Vision (ICCV). pp. 24068–24077 (2025)
16. Liew, S.H., Choo, Y.H., Low, Y.F.: Missing values imputation using similarity matching method for brainprint authentication. International Journal of Advanced Computer Science and Applications **9**(10) (2018)
17. Liu, Y.K., Jiang, W.B., Lu, B.L.: Increasing the stability of EEG-based emotion recognition with a variant of neural processes. In: International Joint Conference on Neural Networks (IJCNN). pp. 1–6 (2022)
18. Liu, Y., De Vos, M., Van Huffel, S.: Compressed sensing of multichannel EEG signals: the simultaneous cosparsity and low-rank optimization. IEEE Transactions on Biomedical Engineering **62**(8), 2055–2061 (2015)
19. Manjur, S.M., Hossain, M.B., Constable, P.A., Thompson, D.A., Marmolejo-Ramos, F., et al.: Detecting autism spectrum disorder using spectral analysis of electroretinogram and machine learning: Preliminary results. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). pp. 3435–3438 (2022)
20. Nickerson, P., Baharloo, R., Davoudi, A., Bihorac, A., Rashidi, P.: Comparison of gaussian processes methods to linear methods for imputation of sparse physiological time series. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). pp. 4106–4109 (2018)
21. Rajan, D., Thiagarajan, J.J.: A generative modeling approach to limited channel ECG classification. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). pp. 2571–2574 (2018)

22. Robson, A.G., Frishman, L.J., Grigg, J., Hamilton, R., Jeffrey, B.G., et al.: ISCEV standard for full-field clinical electroretinography (2022 update). Documenta Ophthalmologica **144**(3), 165–177 (2022)
23. Šuštar Habjan, M., Bach, M., van Genderen, M., Li, S., Mizota, A., et al.: ISCEV standard for clinical visual evoked potentials (2025 update). Documenta Ophthalmologica **151**(2), 97–112 (2025)
24. Tan, M., Le, Q.: EfficientNet: Rethinking model scaling for convolutional neural networks. In: International Conference on Machine Learning (ICML). pp. 6105–6114 (2019)
25. Tashiro, Y., Song, J., Song, Y., Ermon, S.: CSDI: Conditional score-based diffusion models for probabilistic time series imputation. Advances in Neural Information Processing Systems (NeurIPS) **34**, 24804–24816 (2021)
26. Usanmaz, C.F., White, M.E., Valancius, M., Wang, M.D.: An expansion for automated cardiac anomaly detection frameworks with multimodal missing data imputation. In: IEEE EMBS International Conference on Biomedical and Health Informatics (BHI). pp. 1–5 (2021)
27. Yao, C., Liu, J., Higashita, R., Han, Z., Liu, J.: Evoked potential-evidenced visual impairment categorization using vision mamba on ERG correspondence features. In: International Conference on Intelligent Medicine and Health. pp. 1–10 (2025)
28. Zeng, Z., Peng, Z., Yang, X., Shen, W.: Missing as masking: Arbitrary cross-modal feature reconstruction for incomplete multimodal brain tumor segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). pp. 424–433 (2024)
29. Zhang, J., Yin, P.: Multivariate time series missing data imputation using recurrent denoising autoencoder. In: IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 760–764 (2019)
30. Zhang, Y., He, N., Yang, J., Li, Y., Wei, D., et al.: mmFormer: Multimodal medical transformer for incomplete multimodal learning of brain tumor segmentation. In: International conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). pp. 107–117 (2022)