NeuroMM 2026: Interictal Epileptiform Discharge Detection and Localization in Multimodal Neuro-Signals

1 Introduction

Understanding the brain through multimodal neurophysiological signals is a fundamental frontier of artificial intelligence. Signals such as EEG, ECG, EMG, and synchronized behavioral video encode complex, dynamic, and partially observable neural processes, posing unique challenges for trustworthy and clinically actionable AI[1, 3, 4]. Epilepsy, affecting over 70 million people worldwide, provides a high-impact testbed for multimodal neuro-intelligence, where Interictal Epileptiform Discharges (IEDs) serve as critical biomarkers for diagnosis and localization[2, 12, 13]. However, accurate IED analysis remains challenging due to biological variability, extreme class imbalance, cross-patient and cross-institution distribution shifts, and pervasive noise in real-world clinical data[5, 7].

Most existing methods treat IED analysis as a unimodal EEG problem, ignoring cross-modal interactions and temporal dependencies that are essential for robust interpretation[15, 16]. This gap between multimodal clinical reality and unimodal AI modeling fundamentally limits generalization, robustness, and clinical applicability. Despite progress in neuro-signal analysis, the field lacks a unified benchmark for multimodal neuro-intelligence. Existing datasets[10, 11] are fragmented across institutions with heterogeneous protocols and annotation standards, preventing systematic integration and fair comparison, and obscuring deeper questions about how AI systems should perceive, reason, and generalize across neurophysiological modalities.

To address these challenges, we propose the Multimodal Neuro-Signal Challenge (NeuroMM 2026), a Grand Challenge that moves beyond EEG-centric benchmarks by establishing a unified framework integrating heterogeneous neurophysiological signals and synchronized behavioral data. NeuroMM reframes epilepsy analysis as a multimodal reasoning problem, enabling systematic investigation of multimodal perception, cross-modal reasoning, and spatio-temporal inference, and positioning neurophysiological data as a new frontier of multimedia research aligned with the vision of ACM Multimedia.

Specifically, NeuroMM 2026 comprises three tracks:

- **Track 1: NMM-Basic-IED.** Focusing on the automated detection of IED events from heterogeneous neurophysiological signals through robust temporal feature extraction.
- **Track 2: NMM-Context-IED.** Focusing on enhancing diagnostic specificity via joint multimodal modeling of physiological signals and synchronized video-based visual context.
- **Track 3: NMM-Source-IED.** Focusing on the precise spatial localization of epileptogenic zones through fine-grained classification of multi-channel IED propagation patterns.

NeuroMM 2026 aims to catalyze interdisciplinary collaboration and position neurophysiological data as a frontier of multimedia research. By providing a unified and rigorous benchmark, we seek to advance trustworthy AI systems capable of robust cross-modal reasoning in complex clinical environments. Through this challenge, we intend to establish a standardized evaluation paradigm that bridges the gap between machine perception and clinical decision-making, ultimately transforming the future of brain health diagnostics.

Table 1: Number of epochs by consciousness status.

State	Total Epochs	IED Epochs	Non-IED Epochs
Wake	11,906	764	11,142
Sleep	13,543	1,752	11,791
Total	25,449	2,516	22,933

2 Dataset

The NeuroMM 2026 Challenge is built upon the **vEpiSet**[1] dataset, a clinically grounded multimodal neuro-signal benchmark designed to support standardized and reproducible research in cross-modal EEG analysis for epilepsy[6]. Unlike existing neuro-signal datasets that are fragmented across institutions and collected under heterogeneous protocols, vEpiSet provides a unified acquisition, annotation, and evaluation framework, establishing a reliable foundation for multimodal Interictal Epileptiform Discharge (IED) detection and reasoning.

All data in vEpiSet were collected under strict medical and ethical regulations. The acquisition protocol follows the Declaration of Helsinki (1964) and its subsequent amendments or equivalent ethical guidelines[9]. Written informed consent for clinical data collection and research use was obtained from all participants or their legal guardians. All personally identifiable information was removed prior to release to ensure privacy protection and ethical compliance.

The dataset contains EEG-centered multimodal long-term recordings from 84 subjects, including 52 epilepsy patients with at least one annotated IED event and 32 normal-control subjects. The age distribution is balanced across multiple life stages, ranging from children to elderly individuals, with adolescents and adults forming the primary population (see Fig. 1). Each subject contributes 20 minutes of continuous EEG recordings, resulting in approximately 28 hours of data in total. All signals were acquired at the Epilepsy Center of Peking Union Medical College Hospital (PUMCH) using a standardized long-term monitoring protocol with a sampling rate of 500 Hz. Recordings follow the international 10–20 electrode system and include additional EEG, ECG, and EMG channels. The data cover resting states, sleep, and daily activities, intentionally preserving real-world artifacts and variability to support ecological validity and robustness evaluation.

Continuous recordings are segmented into 4-second epochs following a unified protocol, yielding 25,449 samples in total (see Table 1). Each segment is annotated with two complementary labels: the presence of IED events and the consciousness state (awake or sleep). This design enables fine-grained analysis of IED patterns across physiological states while reflecting realistic clinical class imbalance, where non-pathological segments significantly outnumber pathological discharges.

All IED annotations and consciousness labels were independently reviewed by at least three senior EEG experts through a multi-stage verification pipeline. Annotation disagreements were resolved via

[1] https://github.com/vepiset/vepiset_dataset

expert consensus to ensure clinical reliability and labeling consistency. As illustrated in Fig. 2, IED events are further categorized into five anatomically grounded subtypes—generalized, frontal, temporal, occipital, and centro-parietal—supporting spatial localization analysis and clinically meaningful interpretation.

Within NeuroMM 2026, vEpiSet is organized into task-oriented splits and evaluation protocols, forming a unified experimental framework for multimodal detection, cross-modal reasoning, and spatial localization of epileptogenic patterns. This design transforms the dataset from a static resource into an active research platform that drives progress in multimodal neuro-intelligence.

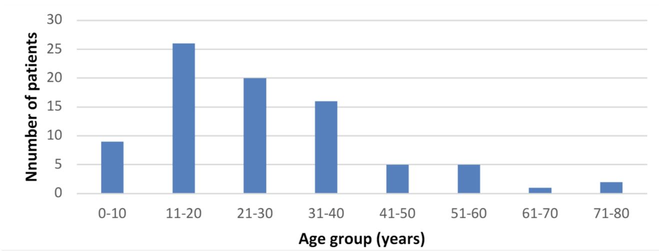

Figure 1: Age distribution of patients in the dataset.

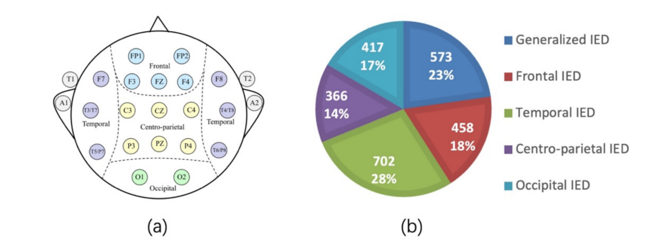

Figure 2: EEG electrodes placement and spatial distribution of interictal epileptiform discharges. (a) Schematic Diagram of EEG electrode placement, including the five spatial regions for classifying IEDs. (b) Percentage and count of each IED type classified by spatial location.

3 Evaluation Metrics

The NeuroMM 2026 Challenge adopts clinically motivated and task-specific evaluation metrics designed to reflect real-world diagnostic requirements, address severe class imbalance, and ensure fair comparison across multimodal learning paradigms. Given the intrinsic rarity of IED events and the critical clinical cost of false positives and missed detections, we emphasize precision-recall–oriented and class-aware metrics rather than accuracy-based measures.

Let TP, FP, TN, and FN denote the numbers of true positives, false positives, true negatives, and false negatives, respectively. Precision and sensitivity (recall) are defined as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{1}$$

NMM-Basic-IED: Robust Multimodal Detection. This track focuses on the fundamental binary detection of IEDs from heterogeneous physiological signals (EEG, ECG, and EMG). In clinical practice, non-IED segments overwhelmingly outnumber pathological events, rendering conventional metrics like accuracy or ROC-AUC insufficient. Therefore, we adopt **Area Under the Precision–Recall Curve (AUPRC)** as the primary metric:

$$\text{AUPRC} = \int_0^1 \text{Precision}(r)\, dr, \tag{2}$$

where r denotes recall. In addition, we report **Precision at Sensitivity = 70% (Precision@Sens=70%)** as an auxiliary metric:

$$\text{Precision@Sens} = 70\% = \text{Precision}\big|_{\text{Sensitivity}=0.7}, \tag{3}$$

reflecting clinically meaningful operating points where high sensitivity is required.

NMM-Context-IED: Vision-Enhanced Robustness. This track evaluates the model's ability to utilize synchronized patient monitoring video as visual context to distinguish true IEDs from motion-induced or environmental artifacts. To protect patient privacy, raw videos are not released; instead, participants utilize pre-extracted visual representations from mainstream models (e.g., CLIP[8], VideoMAE[14]). The task remains binary IED detection based on joint modeling of physiological signals and visual features. Consistent with NMM-Detect, we employ **AUPRC** as the primary metric and **Precision@Sens=70%** as the auxiliary metric, highlighting robustness against artifact-driven false positives.

NMM-Source-IED: Epileptogenic Zone Localization. This track addresses the spatial localization of IED-related brain regions, formulated as a five-class classification problem (generalized, frontal, temporal, occipital, and centro-parietal). For each class c, the F1-score is defined as:

$$\text{F1}_c = \frac{2 \cdot \text{Precision}_c \cdot \text{Recall}_c}{\text{Precision}_c + \text{Recall}_c}. \tag{4}$$

We adopt **Weighted F1-score** as the primary metric:

$$\text{Weighted-F1} = \sum_{c=1}^{C} w_c \cdot \text{F1}_c, \quad w_c = \frac{N_c}{\sum_{k=1}^{C} N_k}, \tag{5}$$

where N_c denotes the number of samples in class c, and C is the total number of classes. As an auxiliary metric, we report the **Macro F1-score**:

$$\text{Macro-F1} = \frac{1}{C} \sum_{c=1}^{C} \text{F1}_c, \tag{6}$$

which emphasizes balanced performance across brain regions.

Overall, the evaluation protocol of NeuroMM 2026 is designed to balance clinical realism, statistical rigor, and multimodal fairness. By aligning metrics with real-world diagnostic constraints and multimodal reasoning objectives, NeuroMM provides a unified and reliable benchmark for advancing multimodal neuro-signal intelligence.

4 Schedule

- **20 April, 2026**: Data, baseline paper & code available
- **20 June, 2026**: Results submission start
- **1 July, 2026**: Results submission deadline
- **12 July, 2026**: Deadline for paper submission
- **5 August, 2026**: Paper acceptance notification
- **19 August, 2026**: Deadline for camera-ready papers

5 People

5.1 Organizer

Qi Tian is the Chief AI Scientist at Huawei CBG and the Director of Guangdong Laboratory of Artificial Intelligence and Digital Economy (SZ). Prior to joining Huawei, he was a Full Professor at the University of Texas at San Antonio, specializing in computer vision, multimedia content analysis, and machine learning. An IEEE Fellow and recipient of the 2021 Wu Wenjun Outstanding Contribution Award in AI, Dr. Tian has a prolific publication record with over 760 papers, including notable awards for best papers at major conferences. His significant contributions to AI in the industry include leading the development of the Pangu series of large-scale pre-trained models, notably Pangu-Weather, which was featured in Nature and ranked as a top scientific advance by the NSFC in 2023.

Fei Ma is currently a researcher at Guangdong Laboratory of Artificial Intelligence and Digital Economy (SZ). Before that, he received the B.S. degree in Communication Engineering from University of Electronic Science and Technology of China in 2017 and the Ph.D. degree in Information and Communication Engineering from Tsinghua University in 2022. So far, he has published over 40 papers in top-tier journals such as IEEE TPAMI, IEEE TMC, IEEE TAFFC, and IEEE TIE, as well as in prestigious conferences including NeurIPS, ACL, AAAI, ACM MM, and IJCAI. His research interests include generative AI and multimodal learning.

Philippe Fournier-Viger is a distinguished professor at Shenzhen University (China), and a national talent in China. He has published 415 research papers related to data mining algorithms for complex data (sequences, graphs), intelligent systems and applications in bioinformatics, which have received more than 18,000 citations (H-Index 66 - Google Scholar). He is the founder of the popular SPMF data mining library, offering more than 250 algorithms to find patterns in data, cited in more than 1,000 research papers. He has extensive editorial experience. He was former associate editor-in-chief of the Applied Intelligence journal (Springer), where he handled over 500 papers. He has also been guest editors of several successful special issues in journals such as EAAI, BDMA, Expert Systems, WCMC, DSPR, INF and IEEE Access. He has been keynote speaker for over 50 international conferences and co-edited four books for Springer. He received the most influential paper award at PAKDD 2024 and the best paper award at five international conferences. He has successfully organized numerous workshops at the PAKDD, PKDD, ICDM, DASFAA, DSAA and KDD conferences. He appears in the top 0.3% of researchers for scientific influence in the Stanford list, and is a Elsevier «Highly Cited Chinese Researcher» (2022).

Zitong Yu received the Ph.D. degree in Computer Science and Engineering from the University of Oulu, Finland, in 2022. He is currently an Associate Professor at Great Bay University, China. He was a postdoctoral researcher at ROSE Lab, Nanyang Technological University, and a visiting scholar at TVG, University of Oxford, from July 2021 to November 2021. His research interests include human-centric computer vision and biometric security. He was a recipient of the IAPR Best Student Paper Award, the IEEE Finland Section Best Student Conference Paper Award 2020, and the second prize of the IEEE Finland Joint Chapter SP/CAS Best Paper Award (2022). He was also recognized in the Stanford World's Top 2% Scientists List (2023–2024).

Laizhong Cui received the B.S. degree from Jilin University, Changchun, China, in 2007, and the Ph.D. degree in computer science and technology from Tsinghua University, Beijing, China, in 2012. He is currently a professor with the College of Computer Science and Software Engineering, Shenzhen University, China. His research interests include future Internet architecture and protocols, edge computing, multimedia systems and applications, blockchain, Internet of Things, cloud computing, and federated learning. He has led more than ten research projects, including the National Key Research and Development Plan of China, the National Natural Science Foundation of China, the Guangdong Natural Science Foundation of China, and the Shenzhen Basic Research Plan. He has published more than 100 papers in leading journals and conferences, including IEEE Journal on Selected Areas in Communications, IEEE Transactions on Computers, IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Knowledge and Data Engineering, IEEE Transactions on Multimedia, IEEE Transactions on Mobile Computing, IEEE Internet of Things Journal, IEEE Transactions on Industrial Informatics, IEEE Transactions on Network and Service Management, ACM Transactions on Internet Technology, IEEE Network, IEEE INFOCOM, ACM MM, IEEE ICNP, and IEEE ICDCS.

5.2 Data Chair

Nan Lin received the B.S. degree in clinical medicine from Tsinghua University in 2009, and the M.D. degree from Peking Union Medical College in 2014. She is currently an associate chief physician at Peking Union Medical College Hospital, Beijing, China. She has specialized in EEG and epilepsy for over 10 years, with expertise in clinical diagnosis and treatment of epilepsy, scalp EEG interpretation, and SEEG interpretation for drug-resistant epilepsy preoperative evaluation. She has published multiple papers in peer-reviewed journals, including BMC Medicine, Neural Networks, Scientific Data, Epilepsy & Behavior, and BMC Neurology. Her research interests include EEG quantitative analysis, deep learning-based automated EEG abnormal pattern detection, and brain-computer interfaces.

Zheng Lian is an Associate Professor at the State Key Laboratory of Autonomous Intelligent Unmanned Systems, Tongji University. He received his Ph.D degree from CASIA, China, in 2021. His research interests primarily center on human-centric AI and affective computing. He co-organized a series of challenges and workshops (MER@IJCAI, MRAC@ACM Multimedia, MEIJU@ICASSP), established benchmark (MER-Bench) and toolbox (MERTools), and proposed new tasks to enhance accuracy and reliability (EMER, OV-MER, AffectGPT, AffectGPT-R1, EmoPrefer). He (co-)authored more than 100 publications in journals, patents, and conference proceedings, including IEEE TPAMI, NeurIPS, ICML, ICLR, IEEE TNNLS, IEEE TASLP, and IEEE TAFFC, leading to >4,400 citations (h-index: 35). He also serves as Associate Editor of IEEE TAFFC and IEEE TASLP, Area Editor of Information Fusion, Area Chair of ACM Multimedia 2025 and ACL ARR 2025, and ACM Multimedia 2026 Dataset Co-Chair.

Haibo He received the Ph.D. degree in Electronic Information Engineering from Tsinghua University. He is currently Deputy General Manager of the Technology Department at NetEase Media and Co-Chair of the NetEase Technology Committee. He is also a member of the CCF CTO Club. Previously, he served as Co-Founder and Chief Technology Officer of Gaifan Media and as Director of NetEase Huati. In 2025, he was selected as the project lead for an AI-enabled medical device innovation initiative under China's Ministry of Industry and Information Technology. His research interests include visual information acquisition and processing, computer vision, natural language processing, and machine learning.

Zebang Cheng is currently a Ph.D. candidate at Shenzhen University, China. His research focuses on affective computing, with particular emphasis on multimodal large models for emotion recognition and reasoning. He has participated in multiple research projects and published several papers in top international venues such as NeurIPS and ACM Multimedia. He also serves as a reviewer for leading journals including IEEE Transactions on Affective Computing, IEEE Transactions on Multimedia, and Pattern Recognition. His work aims to deepen the understanding of the cognitive mechanisms underlying human emotions and to advance the development of reliable and empathetic human–AI interaction.

5.3 Program Committee

Yisu Dong (NetEase Media Technology), Zi Liang (NNetEase Media Technology), Lian Li (NetEase Media Technology), Peng Hu (NetEase Media Technology), Weifang Gao (Department of Neurology, Peking Union Medical College Hospital), Heyang Sun (Department of Neurology, Peking Union Medical College Hospital), Qiang Lu (Department of Neurology, Peking Union Medical College Hospital), Hongbo Xu (Guangming Laboratory), Minghui Li (Guangming Laboratory).

6 Commitment

If our proposal is approved, we formally commit to maintaining a dedicated Grand Challenge website for at least the next three years, hosting all relevant information, benchmark resources, and competition tasks. We certify that the provided neuro-signal data was collected under the Declaration of Helsinki with full informed consent and rigorous de-identification to ensure patient privacy; all materials will be used exclusively for non-commercial academic research.

Furthermore, we pledge to collaborate closely with the ACM Multimedia organizers to promote the challenge, ensuring a fair, open, and transparent competition for all participants. Building on the established multi-track framework, we commit to sustaining this initiative annually by continuously expanding high-quality annotations and introducing increasingly diverse multimodal neuro-intelligence tasks to foster long-term growth within the research community.

For any inquiries regarding the challenge, please contact: **Fei Ma** (mafei@gml.ac.cn) and **Nan Lin** (lin-n06@163.com).

References

[1] Alexander Craik, Yongtian He, and Jose L Contreras-Vidal. 2019. Deep learning for electroencephalogram (EEG) classification tasks: a review. *Journal of neural engineering* 16, 3 (2019), 031001.

[2] E Dessevres, M Valderrama, and M Le Van Quyen. 2025. Artificial intelligence for the detection of interictal epileptiform discharges in EEG signals. *Revue Neurologique* (2025).

[3] P Grosse, MJ Cassidy, and P Brown. 2002. EEG–EMG, MEG–EMG and EMG–EMG frequency analysis: physiological principles and clinical applications. *Clinical Neurophysiology* 113, 10 (2002), 1523–1531.

[4] Espen Hagen, Solveig Næss, Torbjørn V Ness, and Gaute T Einevoll. 2018. Multimodal modeling of neural network activity: computing LFP, ECoG, EEG, and MEG signals with LFPy 2.0. *Frontiers in neuroinformatics* 12 (2018), 92.

[5] Nan Lin, Weifang Gao, Lian Li, Junhui Chen, Zi Liang, Gonglin Yuan, Heyang Sun, Qing Liu, Jianhua Chen, Liri Jin, et al. 2024. vEpiNet: A multimodal interictal epileptiform discharge detection method based on video and electroencephalogram data. *Neural Networks* 175 (2024), 106319.

[6] Nan Lin, Mengxuan Zheng, Lian Li, Peng Hu, Weifang Gao, Heyang Sun, Chang Xu, Gonglin Yuan, Zi Liang, Yisu Dong, et al. 2025. An EEG dataset for interictal epileptiform discharge with spatial distribution information. *Scientific Data* 12, 1 (2025), 229.

[7] Iyad Obeid and Joseph Picone. 2016. The temple university hospital EEG data corpus. *Frontiers in neuroscience* 10 (2016), 196.

[8] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. 2021. Learning transferable visual models from natural language supervision. In *International conference on machine learning*. PmLR, 8748–8763.

[9] Povl Riis. 2003. Thirty years of bioethics: the Helsinki Declaration 1964-2003. *New Review of Bioethics* 1, 1 (2003), 15–25.

[10] Vinit Shah, Eva Von Weltin, Silvia Lopez, James Riley McHugh, Lillian Veloso, Meysam Golmohammadi, Iyad Obeid, and Joseph Picone. 2018. The temple university hospital seizure detection corpus. *Frontiers in neuroinformatics* 12 (2018), 83.

[11] Ali Hossam Shoeb. 2009. *Application of machine learning to epileptic seizure onset detection and treatment.* Ph. D. Dissertation. Massachusetts Institute of Technology.

[12] Anuradha Singh and Stephen Trevick. 2016. The epidemiology of global epilepsy. *Neurologic clinics* 34, 4 (2016), 837–847.

[13] Samuel B Tomlinson, Benjamin C Kennedy, and Eric D Marsh. 2025. Co-activation of interictal epileptiform discharges localizes seizure onset zone and fluctuates with brain state. *Brain Communications* 7, 2 (2025), fcaf127.

[14] Zhan Tong, Yibing Song, Jue Wang, and Limin Wang. 2022. Videomae: Masked autoencoders are data-efficient learners for self-supervised video pre-training. *Advances in neural information processing systems* 35 (2022), 10078–10093.

[15] Jiquan Wang, Sha Zhao, Zhiling Luo, Yangxuan Zhou, Haiteng Jiang, Shijian Li, Tao Li, and Gang Pan. 2024. Cbramod: A criss-cross brain foundation model for eeg decoding. *arXiv preprint arXiv:2412.07236* (2024).

[16] Xujia Wang, Xuhui Liu, Xi Liu, Qian Si, Zhaoliang Xu, Yang Li, and Xiantong Zhen. 2025. Eeg-dino: Learning eeg foundation models via hierarchical self-distillation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 196–205.