scientific data

Check for updates

OPEN

DATA DESCRIPTOR

An EEG dataset for interictal epileptiform discharge with spatial distribution information

Nan Lin[1], Mengxuan Zheng[2], Lian Li[2], Peng Hu[2], Weifang Gao[1], Heyang Sun[1], Chang Xu[2], Gonglin Yuan[2], Zi Liang[2], Yisu Dong[2], Haibo He[2], Liying Cui[1] ✉ & Qiang Lu[1] ✉

Interictal epileptiform discharge (IED) and its spatial distribution are critical for the diagnosis, classification, and treatment of epilepsy. Existing publicly available datasets suffer from limitations such as insufficient data amount and lack of spatial distribution information. In this paper, we present a comprehensive EEG dataset containing annotated interictal epileptic data from 84 patients, each contributing 20 minutes of continuous raw EEG recordings, totaling 28 hours. IEDs and states of consciousness (wake/sleep) were meticulously annotated by at least three EEG experts. The IEDs were categorized into five types based on occurrence regions: generalized, frontal, temporal, occipital, and centro-parietal. The dataset includes 2,516 IED epochs and 22,933 non-IED epochs, each 4 seconds long. We developed and validated a VGG-based model for IED detection using this dataset, achieving improved performance with the inclusion of consciousness and/or spatial distribution information. Additionally, our dataset serves as a reliable test set for evaluating and comparing existing IED detection models.

Background & Summary

Epilepsy is a common, chronic, and complex group of neurological disorders, affecting more than 50 million individuals across the globe[1]. Electroencephalography (EEG), which measures neuronal activities, serves as a primary diagnostic tool for epilepsy. Interictal epileptiform discharge (IED) is typically found in patients with epilepsy[2], exhibiting diverse appearances depending on the type of epilepsy[3]. In current clinical practice, visual analysis remains the gold standard for EEG interpretation[2]. However, manual detection of IEDs from EEG recordings requires considerable skill and is time-consuming. This motivates the development of automated IED detection algorithms, reducing the time and resources spent on visual analysis, as well as the misdiagnosis rate[4].

Deep learning (DL)-based methods have been widely developed and employed on EEG datasets for seizure and IED detection, achieving promising results[4–7]. However, most existing DL models are trained and validated on private datasets[5,6,8,9], and their performance relies heavily on the quality, quantity, and distribution of the EEG signals in these datasets. The use of private datasets restricts the reproducibility and independent verification of the results, making it challenging for other researchers to benchmark and compare different models. Additionally, the lack of access to these datasets impedes further improvements in model accuracy and generalizability.

The availability of publicly accessible annotated IED data is crucial for advancing DL methods. However, most existing publicly available EEG datasets[10–12] suffer from small sample sizes, typically ranging from 6 to 30 subjects, and often lack IED annotations, as highlighted by Palak Handa *et al.*[13]. Given the significant variability and individual differences in EEG signals, DL models based on limited EEG datasets may encounter challenges of low accuracy and generalizability. Therefore, the availability of large, publicly accessible annotated IED datasets is essential for advancing DL methods.

The Temple University EEG corpus (TUH-EEG Corpus) is a popular public dataset, containing 19,057 annotated IEDs and classifying the EEG events into six classes, including spike and/or sharp waves and generalized periodic epileptiform discharges[14]. However, it lacks spatial information on the EEG recordings. This spatial information is crucial, given that epilepsy is a chronic brain disease causing recurrent spontaneous seizures that necessitate anti-seizure medication (ASM). Such spatial distribution information can assist in the classification

[1]Department of Neurology, Peking Union Medical College Hospital, Beijing, 100730, China. [2]NetEase Media Technology Co., Ltd, Beijing, 100084, China. ✉e-mail: pumchcuily@sina.com; luqiang@pumch.cn

of epilepsy types and strategizing appropriate ASM. Additionally, this information aids in localizing the epileptic focus, a crucial element in presurgical assessments[15]. To the best of our knowledge, the dataset proposed in this paper is the first publicly available dataset containing the spatial distribution information of IEDs. Furthermore, due to the differing physiological activities and EEG signals during wakefulness and sleep — factors that can interfere with and influence IED judgment — our dataset also includes annotations detailing sleep-wake states. This comprehensive dataset is expected to meet the need for detailed spatial and sleep-wake state annotations. A dataset enriched with both temporal and spatial information is crucial not only for enhancing the accuracy and efficiency of epilepsy diagnosis and management but also for advancing the development and validation of DL models in the field of neurology.

Many researchers have employed cross-validation to evaluate the performance of IED detection models[5,6,16]. In this scenario, the evaluation dataset contains a high ratio of IED recordings, which is different from the real-world EEG recordings that contain much more normal EEG data. Such an imbalanced data distribution may result in inaccurate evaluations. Additionally, artifacts from the environment and body movements are unavoidable, potentially causing a higher false positive rate. Our previous study has demonstrated that using cross-validation can overestimate a model's performance[17]. There is a need for a dataset comprising raw continuous EEG data that reflects real-world clinical scenarios to assist in fair evaluations and comparisons for automatic IED detection models.

To address these needs, we present an EEG dataset containing recordings in NPY format from 84 Chinese epilepsy patients. Each patient contributed 20 minutes of raw continuous EEG data, resulting in a total of 28 hours of recordings. Among these, 52 recordings contain at least one IED, meticulously annotated and verified by at least three EEG experts. Information regarding spatial distribution (categorized into generalized, frontal, temporal, occipital, and centro-parietal regions) and consciousness state (wake/sleep) at the time of IED occurrence is also provided. Additionally, EEG data from the remaining 32 subjects with normal EEG reports underwent the same segmentation process to serve as negative samples. This process resulted in 2,516 IED-containing 4-second EEG epochs (9.9%) and 22,933 non-IED epochs (90.1%). We trained and validated a VGG[18]-based IED detection model using our dataset to demonstrate its robustness and effectiveness in establishing artificial intelligence models for epilepsy diagnosis and management. VGG is a commonly used classical convolutional neural network architecture for its excellent performance and adaptability in various tasks. Furthermore, we used this dataset as a test set in a simulated real-world environment, employing a previously published IED automatic detection model (trained independently from our dataset)[17] to demonstrate its suitability for simulating real-world testing scenarios and facilitating comparisons among different IED detection models.

In this paper, we propose a comprehensive dataset designed for the training and evaluation of DL-based models for automatic IED detection. This dataset includes a sufficient number of samples, detailed spatial categorization of epileptiform discharges during interictal periods, sleep-wake annotations, and simulated real-world testing data. It provides EEG data from the Chinese population, increasing the diversity of public IED data. The EEG recordings are annotated with IED labels, categorized into five spatial locations (generalized, frontal, temporal, occipital, and centro-parietal regions), and distinguish between sleep and wake states. The dataset is publicly available at Figshare[19].

This database supports various studies, including:

1. Algorithm development for automated IED detection;
2. Automated classification of IEDs for epilepsy types;
3. Differentiation of patient consciousness states to improve detection accuracy;
4. Testing in simulated real-world clinical scenarios for existing models.

Methods

The study was conducted in accordance with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Patients or patients' parents (for patients under the age of 18) had provided written informed assent for the Collection and Application of Clinical Sample and Medical Data at the time of admission, including publish and public of clinical data without direct identifiable information, as certified and approved by the ethics committee of PUMCH upon their hospital admission (NO. I-22PJ169). To protect participant information, we anonymise all identifiable information, including names and ID numbers. However, demographic data, such as age and gender, which can affect the EEG data, are included in the dataset.

Participants. The dataset comprises 52 epilepsy patients (33 males and 19 females, median age 24, range 5–72) who exhibited at least one IED, and 32 patients (24 males and 8 females, median age 30, range 7–77) with normal EEG recordings. Fourteen of the 32 patients were diagnosed with epilepsy. The age distribution is illustrated in Fig. 1. The data were collected from the Epilepsy Center at Peking Union Medical College Hospital (PUMCH) between March 2023 and March 2024. Fifty-nine epilepsy patients received antiepileptic drugs, while 25 had not received any ASM.

EEG recording. Each patient underwent a long-term ($\geq$ 2 hours) EEG recording (EEG-1200C, Nihon Kohden, Tokyo, Japan) at PUMCH. The 20-minute continuous EEG recordings were extracted by an epilepsy specialist (NL) and encompassed various clinical scenarios, such as resting with eyes closed, resting with eyes opened, sleeping, and daily activities (eating, limb movements, etc.). It is noted that since the artifacts from daily activities are highly variable, and IED detection models should distinguish IEDs from all types of artifacts, patients' positions are not annotated during EEG monitoring and collection. Although the presence of artifacts

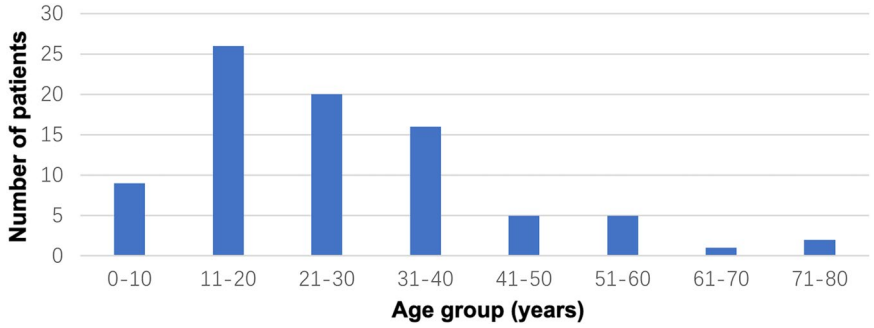

Fig. 1 Age distribution of patients in the dataset.

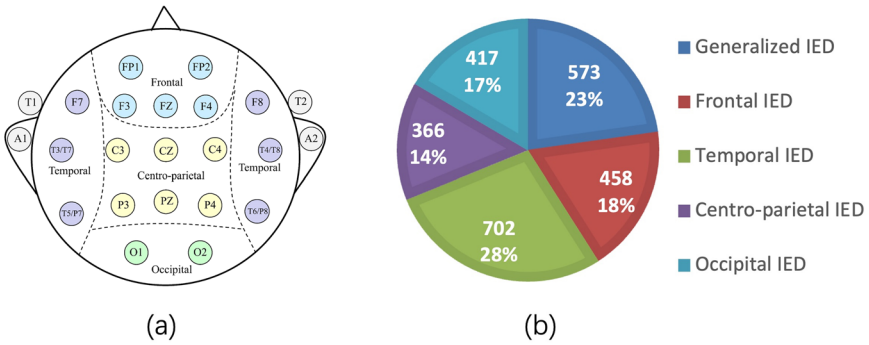

Fig. 2 EEG electrodes placement and spatial distribution of interictal epileptiform discharges. (**a**) Schematic Diagram of EEG electrode placement, including the five spatial regions for classifying IEDs. (**b**) Percentage and count of each IED type classified by spatial location.

might challenge IED detection, it replicates real clinical settings and accurately reflects the performance of the model in clinical practice. The 20-minute EEG recordings from the 52 patients with epilepsy each contain at least one IED.

As depicted in Fig. 2(a), the EEG recordings include 19 standard 10–20 system electrodes (Fp1, F3, C3, P3, O1, F7, T3, T5, Fz, Cz, Pz, Fp2, F4, C4, P4, O2, F8, T4, T6), supplemented with T1, T2, A1, A2, two electrocardiographic (ECG) electrodes (RA and LA, placed under the clavicles), and four electromyographic (EMG) electrodes placed on deltoid muscles (two on each side), all sampled at 500 Hz. The EEG data are 16 bits.

The EEG data were uniformly preprocessed by applying a 0.1–70 Hz filter using the SciPy toolbox for Python 3.9, thereby minimizing interference from high-frequency muscle activity and baseline drift. A 50 Hz notch filter was applied to eliminate power line noise. This preprocessing method ensures that the data are clean and standardized, providing a reliable foundation for further analysis and application in various research and clinical scenarios.

EEG annotation. The initial identification of IEDs was performed by an EEG expert (HS) during routine clinical assessments based on the recognized definitions[20]: IEDs are transient waveforms that emerge from the cortical regions of the brain, characterized by their unique spike or sharp wave morphology, and typically lasting between 20 to 200 milliseconds. After extracting the targeted 20-minute EEG segments, an epilepsy specialist (NL) and an experienced EEG technologist (WG), both qualified and experienced in EEG interpretation, independently reassessed the IED annotations, mapped the IED distribution, and categorized the EEG data into wake and sleep periods. In cases of disagreement, a third EEG expert (QL) made the final decision. The durations of most spikes/sharps ranged from 40 to 240 ms, and their amplitudes ranged from 16 to 184 μV. The duration and amplitude of spike/sharp and wave complexes mostly ranged from 160 to 660 ms and from 26 to 364 μV, respectively.

The raw EEG records were annotated by at least three experts using the Neuro Workbench software (NIHON KOHDEN Corporation, JP) with an annotation temporal resolution of 0.002 seconds. The raw EEG data and the annotated tags are stored together in MAT files[19]. The annotations include the following points:

(1) IEDs were annotated using an exclamation mark (!) at the corresponding spike/sharp location. For instances of continuous or periodic (>1 Hz) discharges, the start and end of IEDs were denoted with '! start' and '! end', respectively (Fig. 3(a,b,e)).
(2) Annotations of consciousness state were made regarding the patient's status, using the terms 'waking' and 'sleeping'. The first annotation to denote the state of consciousness was placed at the beginning of each EEG recording (Fig. 3(c)). Any subsequent changes in the patient's state were recorded with the corresponding

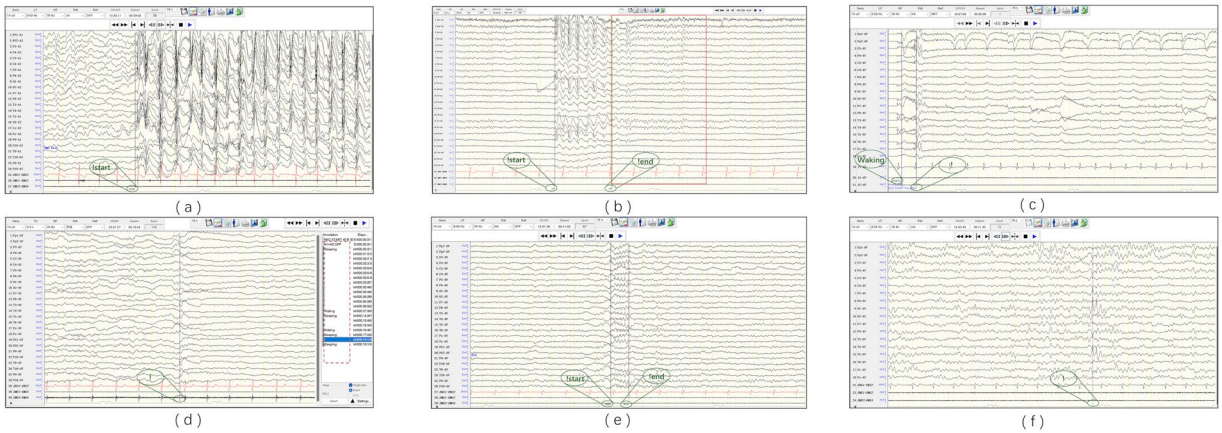

Fig. 3 EEG annotations and spatial distribution classification of interictal epileptiform discharges (IEDs). (**a**-**b**) Generalized IEDs. Continuous discharges are labeled from '!start' to '!end'. (**b**) An example of truncated IEDs. The red square represents a 4-second segment containing the label '!end' but no IED. (**c**) Left frontal IED and consciousness state (waking) are labeled at the beginning of the EEG recording. (**d**) Left temporal IED. The list on the right displays all annotations in this EEG data, including consciousness states and IEDs. (**e**) Centro-parietal IED. The patient presented with bilateral limb myoclonus. (**f**) Occipital IED.

Consciousness 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

Table 1. Number of epochs by consciousness status.

annotations. Taking the Fig. 3(d) as an example, the first annotation in the given EEG data was 'sleeping', which indicates that the patient began in a sleep state. If the patient's state transitioned into an awake state, a 'waking' annotation was added. The period between these two annotations was classified as a sleep state. The division between wakefulness and sleep was conducted following the American Academy of Sleep Medicine manual for sleep scoring[21]. Due to the absence of the electrooculogram data, sleep stages were not further categorized.

The raw EEG data were subsequently segmented into 4-second epochs based on the annotation to facilitate analysis and validation in deep learning. For each epoch labelled with IEDs, the epilepsy specialist (NL) re-evaluated the data to confirm the presence of IEDs. If an IED is truncated, the spike could be included in the preceding epoch, while the following wave with annotation (such as '!end') appears in the next epoch (Fig. 3(b)). Three truncated positive samples were removed, resulting in a total of 2,516 epochs with IEDs (positive samples) and 22,933 non-IED epochs (negative samples). The IEDs were classified into five types based on their occurrence regions within the 4-second epochs, and the segmentation criteria are shown in Fig. 2[22]: (1) generalized IED, including bilateral synchronous IED predominantly over the anterior or posterior head; (2) frontal IED; (3) temporal IED; (4) centro-parietal IED, including the discharges primarily occurring over the central/parietal regions, corresponding to the anterior or posterior central gyrus; (5) occipital IED, including the epileptiform discharges over the temporal-parieto-occipital region (the posterior head region), but predominantly over the occipital lobe (Table 1). Examples of these five discharge types are shown in Fig. 3(a-f). Furthermore, patients could exhibit multiple types of IEDs in various scenarios: (1) multiple focal IEDs; (2) generalized IEDs can appear fragmented and have focal features, which are categorized as focal types[20]; (3) focal IEDs can spread widely with generalized features, which categorized as generalized IEDs[23].

Data Records

The dataset provided in this paper is available at Figshare[19]. The structure of the dataset is illustrated in Fig. 4. The MAT-Files folder contains 84 MAT files and a base_info.csv file that records the basic information (gender, age) of the patients. Each MAT file encompasses 20 minutes of raw EEG record from a patient, along with two types of event annotations: timestamps for IED discharges and markers for wakefulness and sleep. The other folders (excluding the MAT-files folder) contain 4-second NPY EEG data segments derived from the MAT files based on event identification. There are 2,516 epochs (positive samples) containing IEDs, categorized into different folders by brain spatial location, and 22,933 negative samples for all patients. The remaining folders include information on whether IEDs occurred and their five categorical spatial locations. The NPY file naming convention, exemplified by "DA00102S170000_172000_500_5", is as follows. Here, "DA00102S" signifies the original MAT filename, "170000" indicates the starting timestamp of the segment, "172000" marks the ending

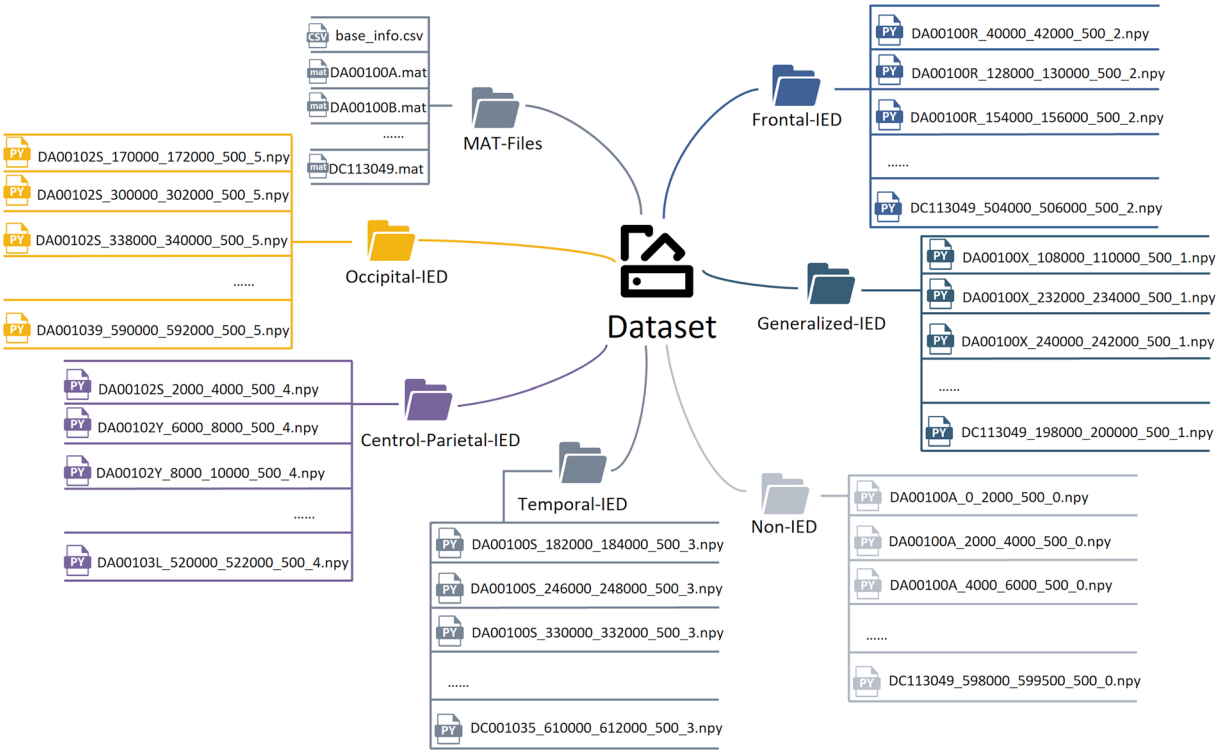

Fig. 4 The structure of the dataset.

	VGG_{EEG}		VGG_{sleep}		vEpiNet	
Sensitivity	Precision	Specificity	Precision	Specificity	Precision	Specificity
50%	96.1%	99.8%	96.5%	99.8%	93.9%	99.6%
55%	94.7%	99.7%	95.1%	99.7%	91.9%	99.5%
60%	94.0%	99.6%	91.6%	99.4%	88.8%	99.2%
65%	91.9%	99.4%	90.4%	99.3%	84.7%	98.7%
70%	87.8%	99.0%	86.2%	98.8%	79.0%	98.0%
75%	82.3%	98.3%	83.1%	98.4%	73.1%	97.0%
80%	74.2%	97.0%	76.3%	97.3%	63.6%	95.0%
85%	63.4%	94.8%	67.1%	95.5%	53.8%	92.0%
90%	50.4%	90.5%	60.0%	93.6%	41.1%	85.9%
95%	29.5%	75.8%	49.4%	89.6%	24.5%	67.9%

Table 2. Sensitivity, specificity, and precision of the vEpiNet and VGG models.

timestamp, "500" represents the sampling rate per second, and "5" designates the category to which the segment belongs. The categories are encoded as follows: 0 stands for non-IED, 1 for generalized IED, 2 for frontal IEDs, 3 for temporal IED, 4 for centro-parietal IED, and 5 for occipital IED. Table 1 and Fig. 2(b) display the number of epochs under each type of label in the segmented dataset.

Technical Validation

We trained a VGG network[18], a widely employed architecture, for IED detection, demonstrating the effectiveness of the proposed dataset. This model is named VGG_{EEG}. The 4-second EEG segments were partitioned into a training subset (18,335 non-IED epochs and 2,024 IED epochs) and a test subset (4,589 non-IED epochs and 492 IED epochs) in an 8:2 ratio. The results show that at a sensitivity of 80%, the precision and specificity of IED detection are 74.2% and 97%, respectively (Table 2). These promising findings indicate that the proposed dataset provides sufficient data for training DL-based model.

Factors influencing IED detection differ between wakefulness and sleep periods. During wakefulness periods, muscle artifacts and movements from daily activities are the primary interferences affecting detection. Conversely, during sleep, physiological sleep patterns such as vertex sharp waves and K complexes predominantly impact detection, rather than movement artifacts. Therefore, the state of consciousness (wake/sleep) is incorporated as an independent feature in the IED detection model. To increase the weight of consciousness state in the model, it is transformed into a 24-dimensional vector using a Multi-Layer Perceptron (MLP) and

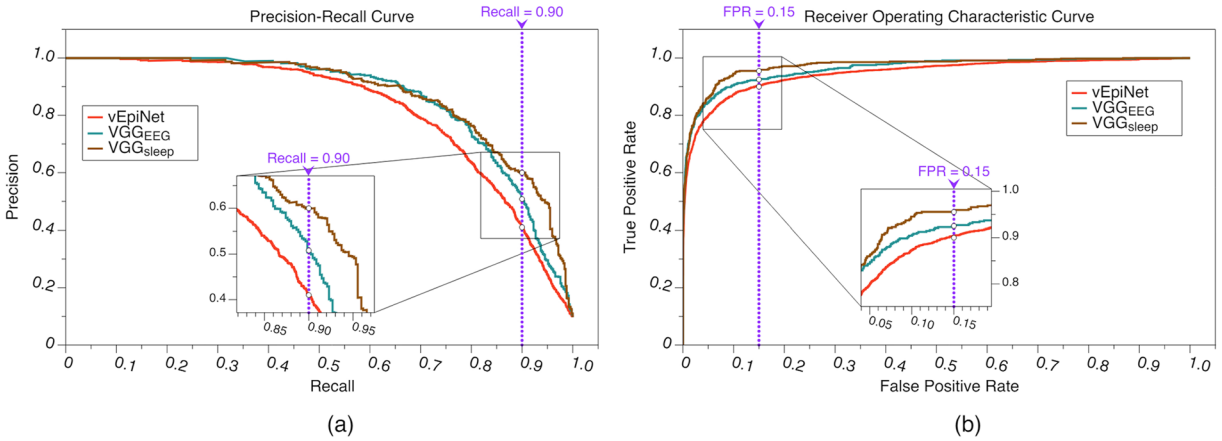

Fig. 5 Validation results of different models. (**a**) Precision-Recall (PR) curves of the VGG_{EEG}, VGG_{sleep}, and vEpiNet models. (**b**) Receiver operating characteristic (ROC) curves of the VGG_{EEG}, VGG_{sleep}, and vEpiNet models.

Predict Label								
		Generalized IED	**Frontal IED**	**Temporal IED**	**Centro-parietal IED**	**Occipital IED**	**Precision**	**Sensitivity**
True Label	**Generalized IED**	504	48	8	5	8	0.913	0.880
	Frontal IED	40	329	68	8	13	0.760	0.718
	Temporal IED	7	41	631	5	18	0.812	0.899
	Centro-parietal IED	0	5	13	348	0	0.935	0.951
	Occipital IED	1	10	57	6	343	0.898	0.823

Table 3. The confusion matrix of 5-fold cross-validation of IED classification based on spatial distribution.

subsequently integrated with EEG features. This model is termed VGG_{sleep}. The results indicate that VGG_{sleep} achieves an area under the precision-recall curve (AUPRC) of 0.8718, surprising VGG_{EEG}'s 0.8585. At a sensitivity of 90.0%, precision is 60.0% for VGG_{sleep} compared to 50.4% for VGG_{EEG}. The improvement in precision is notably pronounced at higher sensitivity (Shown in Table 2 and Fig. 5), indicating that the information on consciousness states provided in our dataset can assist in enhancing the performance of DL-based models.

To demonstrate that our dataset can serve as a test set simulating real clinical scenarios for existing IED detection models, we employed the proposed dataset to test our previously published EEG detection model, vEpiNet[17]. Please note that the training data for vEpiNet is independent of the dataset provided in this paper. The results show a precision of 63.6% and a specificity of 95.0% at an 80.0% sensitivity (Table 2), consistent with the prospective test results of vEpiNet on 50 patients, confirming that the data distribution in the proposed dataset is similar to that of a real clinical environment.

Table 3 exhibits a confusion matrix derived from five-fold cross-validation, tailored for the classification of five types of IEDs. The VGG model has demonstrated remarkable performance, achieving over 70% in both precision and sensitivity. This result further validates the high applicability and effectiveness of our dataset in training and classifying IED types based on spatial distribution.

In summary, this paper proposes a Chinese EEG dataset containing 84 subjects and 2,516 IED epoch. The efficacy of the dataset for IED detection and classification based on spatial location has been validated. Experimental results further indicate that integrating consciousness states improves model performance.

Usage Notes

We have successfully accessed the MAT files using the MATLAB software (https://www.mathworks.com). We have also successfully imported the MAT files into Python and segmented events into numpy files.

For researchers working on automated algorithms for EEG analysis, additional material can be found on Github (https://github.com/vepiset/vepiset_dataset). These repositories offer guidance on dataset creation, algorithms for detecting interictal epileptiform discharges, and methods for evaluating algorithm outputs.

The dataset comprising real patients exhibiting interictal epileptiform discharges inherently possesses potential limitations, as the recorded signal data can be susceptible to variations stemming from patient movement, electrocardiogram interference, and other contributing factors. Nonetheless, this collection of data serves as a valuable resource, offering multi-dimensional identifiers that enhance understanding within the realm of electroencephalography and foster groundbreaking advancements.

Code availability

The custom code used to access and analyze this dataset is available at https://github.com/vepiset/vepiset_dataset.

Received: 12 July 2024; Accepted: 31 January 2025;
Published online: 07 February 2025

References

1. World Health Organization. Epilepsy: a public health imperative. (2019).
2. Kural, M. A. *et al*. Criteria for defining interictal epileptiform discharges in EEG: A clinical validation study. *Neurology* **94**, e2139–e2147, https://doi.org/10.1212/wnl.0000000000009439 (2020).
3. Tatum, W. O. *et al*. Clinical utility of EEG in diagnosing and monitoring epilepsy in adults. *Clin Neurophysiol* **129**, 1056–1082, https://doi.org/10.1016/j.clinph.2018.01.019 (2018).
4. da Silva Lourenço, C., Tjepkema-Cloostermans, M. C. & van Putten, M. Machine learning for detection of interictal epileptiform discharges. *Clin Neurophysiol* **132**, 1433–1443, https://doi.org/10.1016/j.clinph.2021.02.403 (2021).
5. Fürbass, F. *et al*. An artificial intelligence-based EEG algorithm for detection of epileptiform EEG discharges: Validation against the diagnostic gold standard. *Clin Neurophysiol* **131**, 1174–1179, https://doi.org/10.1016/j.clinph.2020.02.032 (2020).
6. Jing, J. *et al*. Development of Expert-Level Automated Detection of Epileptiform Discharges During Electroencephalogram Interpretation. *JAMA Neurol* **77**, 103–108, https://doi.org/10.1001/jamaneurol.2019.3485 (2020).
7. Nhu, D. *et al*. Automated Interictal Epileptiform Discharge Detection from Scalp EEG Using Scalable Time-series Classification Approaches. *International Journal of Neural Systems* **33**, 2350001 (2023).
8. Thomas, J. *et al*. Automated Adult Epilepsy Diagnostic Tool Based on Interictal Scalp Electroencephalogram Characteristics: A Six-Center Study. *Int J Neural Syst* **31**, 2050074, https://doi.org/10.1142/s0129065720500744 (2021).
9. Nhu, D. *et al*. Deep learning for automated epileptiform discharge detection from scalp EEG: A systematic review. *J Neural Eng* **19**, https://doi.org/10.1088/1741-2552/ac9644 (2022).
10. Cserpan, D. *et al*. Dataset of EEG recordings of pediatric patients with epilepsy based on the 10-20 system. *OpenNeuro* (2023).
11. Detti, P. Siena scalp EEG database. *PhysioNet. doi* **10**, 493 (2020).
12. Nasreddine, W. Epileptic EEG dataset. *Mendeley Data, V1* https://doi.org/10.17632/5pc2j46cbc.1 (2021).
13. Handa, P., Mathur, M. & Goel, N. Open and free EEG datasets for epilepsy diagnosis. *arXiv preprint arXiv:2108.01030* (2021).
14. Obeid, I. & Picone, J. The Temple University Hospital EEG Data Corpus. *Front Neurosci* **10**, 196, https://doi.org/10.3389/fnins.2016.00196 (2016).
15. Chvojka, J. *et al*. The role of interictal discharges in ictogenesis - A dynamical perspective. *Epilepsy Behav* **121**, 106591, https://doi.org/10.1016/j.yebeh.2019.106591 (2021).
16. Thomas, J. *et al*. EEG classification via convolutional neural network-based interictal epileptiform event detection. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE 3148-3151 (2018).
17. Lin, N. *et al*. vEpiNet: A multimodal interictal epileptiform discharge detection method based on video and electroencephalogram data. *Neural Netw* **175**, 106319, https://doi.org/10.1016/j.neunet.2024.106319 (2024).
18. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations (ICLR) 1-14* (2015).
19. Lin, N. *et al*. An EEG dataset for interictal epileptiform discharge with spatial distribution information. *figshare* https://doi.org/10.6084/m9.figshare.28069568 (2024).
20. Hirsch, E. *et al*. ILAE definition of the Idiopathic Generalized Epilepsy Syndromes: Position statement by the ILAE Task Force on Nosology and Definitions. *Epilepsia* **63**, 1475–1499, https://doi.org/10.1111/epi.17236 (2022).
21. Berry, R. B. *et al*. The AASM manual for the scoring of sleep and associated events. *Rules, Terminology and Technical Specifications. Darien, Illinois, American Academy of Sleep Medicine* **176**, 7 (2012).
22. Scherg, M., Ille, N., Bornfleth, H. & Berg, P. Advanced tools for digital EEG review: virtual source montages, whole-head mapping, correlation, and phase analysis. *J Clin Neurophysiol* **19**, 91–112, https://doi.org/10.1097/00004691-200203000-00001 (2002).
23. Riney, K. *et al*. International League Against Epilepsy classification and definition of epilepsy syndromes with onset at a variable age: position statement by the ILAE Task Force on Nosology and Definitions. *Epilepsia* **63**, 1443–1474, https://doi.org/10.1111/epi.17240 (2022).

Acknowledgements

This study was supported by National Key Research and Development Project (grant number: 2022YFC2503800), CAMS Innovation Fund for Medical Sciences (CIFMS) (grant number: 2023-I2M-C&T-B-023), National High Level Hospital Clinical Research Funding (grant number: 2022-PUMCH-A-168), Science and Technology Program of the Joint Fund of Scientific Research for the Public Hospitals of Inner Mongolia Academy of Medical Sciences (grant number: 2024GLLH0452).

Author contributions

Lin Nan: Writing – original draft, Formal analysis, Data curation. Zheng Mengxuan: Writing the original draft and formal analysis. Li Lian and Hu Peng: Methodology, Formal analysis. Gao Weifang and Sun Heyang: Data curation. Xu Chang: Writing-original draft. Yuan Gonglin: Writing- review & editing. Liang Zi, Dong Yisu and He Haibo: Methodology. Cui Liying: Writing- review & editing, formal analysis. Lu Qiang: Writing – review & editing, Data curation, Conceptualization.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Additional information

Correspondence and requests for materials should be addressed to L.C. or Q.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.