

ARTICLE OPEN



An interpretable model based on graph learning for diagnosis of Parkinson's disease with voice-related EEG

Shuzhi Zhao^{1,2,3,7}, Guangyan Dai^{1,7}, Jingting Li^{1,7}, Xiaoxia Zhu¹, Xiyan Huang⁴, Yongxue Li¹, Mingdan Tan¹, Lan Wang^{3,5}, Peng Fang^{3,5}, Xi Chen¹, Nan Yan^{3,5} and Hanjun Liu^{1,6}

Parkinson's disease (PD) exhibits significant clinical heterogeneity, presenting challenges in the identification of reliable electroencephalogram (EEG) biomarkers. Machine learning techniques have been integrated with resting-state EEG for PD diagnosis, but their practicality is constrained by the interpretable features and the stochastic nature of resting-state EEG. The present study proposes a novel and interpretable deep learning model, graph signal processing-graph convolutional networks (GSP-GCNs), using event-related EEG data obtained from a specific task involving vocal pitch regulation for PD diagnosis. By incorporating both local and global information from single-hop and multi-hop networks, our proposed GSP-GCNs models achieved an averaged classification accuracy of 90.2%, exhibiting a significant improvement of 9.5% over other deep learning models. Moreover, the interpretability analysis revealed discriminative distributions of large-scale EEG networks and topographic map of microstate MS5 learned by our models, primarily located in the left ventral premotor cortex, superior temporal gyrus, and Broca's area that are implicated in PD-related speech disorders, reflecting our GSP-GCN models' ability to provide interpretable insights identifying distinctive EEG biomarkers from large-scale networks. These findings demonstrate the potential of interpretable deep learning models coupled with voice-related EEG signals for distinguishing PD patients from healthy controls with accuracy and elucidating the underlying neurobiological mechanisms.

npj Digital Medicine (2024)7:3; <https://doi.org/10.1038/s41746-023-00983-9>

INTRODUCTION

Parkinson's disease (PD) is a neurodegenerative disorder that exerts a profound impact on the quality of life for 7–10 million people worldwide^{1,2}. It is characterized by progressive and diverse symptoms that involve both motor and non-motor impairments. However, the pathogenic mechanism of PD remains poorly understood, with only 20% of cases being attributed to specific genetic factors^{3,4}. Therefore, the precise and early diagnosis of PD continues to present considerable challenges, as this holds significance for effective clinical management.

One promising avenue for PD diagnosis lies in the identification of reliable biomarkers across various behavior domains, including handwriting patterns⁵, motor function⁶, gait patterns⁷, and speech characteristics⁸. Of particular interest, resting-state electroencephalography (EEG) has emerged as a potential diagnostic tool for PD diagnosis due to its noninvasiveness, cost-effectiveness, and ability to capture brain activity with high-temporal resolution^{9–11}. Quantitative EEG (QEEG) measures, including power spectral density^{10,11} and spatiotemporal microstates⁹, have been extracted as distinctive features to distinguish PD patients from healthy individuals. More recently, an increasing number of studies on PD diagnosis have shifted towards integrating deep learning techniques with large-scale EEG networks^{12,13}. For example, Oh et al.¹⁴ utilized a thirteen-layer convolutional neural network (CNN) for identifying resting-state EEG data from PD patients, achieving a remarkable classification accuracy (ACC) of 88.25%. Chaturvedi et al.¹⁵ integrated resting-state EEG parameters with least

absolute shrinkage and selection operator (LASSO) and achieved an area under the curve (AUC) of 0.76 in PD diagnosis. These methods, however, rely heavily on their assumption of stationarity and integrability of the EEG signals⁹, which may not be valid given the dynamic nature of PD-related changes in brain activity. Therefore, the need to capture the stable and time-varying patterns, which cannot be adequately addressed by the stochastic resting-state EEG signals, arises as a challenge to extract discriminative features for distinguishing PD patients from healthy individuals.

In contrast to resting-state EEG signals, task-related EEG signals exhibit phase- and time-locked responses to motor and non-motor tasks and their functional networks/connectivity. This aspect offers valuable insights into PD-related alterations in neural activity and extraction of distinct EEG features between PD patients and healthy controls. A particular area of interest lies in motor speech disorders, which affect approximately 90% of PD patients and are considered as one of the premotor symptoms^{16,17}. Previous studies have demonstrated that PD patients are impaired in sensorimotor control of vocal production^{18–22}, which is manifested as enhanced event-related potential (ERP) P2 responses to pitch perturbations in voice auditory feedback^{23–25}. This observation suggests the potential of machine learning-based extraction of salient features from voice-related EEG signals to obtain robust biomarkers for PD diagnosis. Notably, Shi et al.²⁶ developed four deep learning architectures, namely the CNN, the Recurrent Neural Network (RNN) and two hybrid models (2D-CNN-

¹Department of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China. ²School of Biomedical Engineering, Sun Yat-sen University, Shenzhen, China. ³CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. ⁴Department of Rehabilitation Medicine, Zhujiang Hospital, Southern Medical University, Guangzhou, China. ⁵Guangdong-Hong Kong-Macao Joint Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. ⁶Guangdong Provincial Key Laboratory of Brain Function and Disease, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, China. ⁷These authors contributed equally: Shuzhi Zhao, Guangyan Dai, Jingting Li. ✉email: chenxi8@mail.sysu.edu.cn; nan.yan@siat.ac.cn; lhanjun@mail.sysu.edu.cn

RNN and 3D-CNN-RNN), for PD diagnosis based on voice-related EEG signals. Their findings showed that the hybrid models outperformed the conventional models, achieving accuracies of 82.89% (3D-CNN-RNN), 81.13% (2D-CNN-RNN), 80.89% (CNN), and 76.00% (RNN). This study provides preliminary evidence for the feasibility of deep learning models with task-related EEG signals to distinguish PD patients from healthy individuals.

Nevertheless, the complexity of deep learning models has raised concerns regarding their interpretability, particularly in understanding what features they can learn from EEG signals and how they relate to the clinical characteristics of PD patients. Therefore, there is a growing demand for enhancing the interpretability of deep learning methods^{27–29}. One promising interpretable model is the multi-category mutual verification³⁰, which integrates graph learning and neurophysiological information models. The graph convolutional network (GCN), a key component of the graph learning model, has achieved strong performance in diagnosing neurodegenerative diseases using neuroimaging data³¹. The neurophysiological information model involves large-scale neural networks for EEG microstate analysis, which has been successfully used for PD diagnosis with resting-state data⁹. However, no studies have yet combined these two methods to enhance model interpretability in PD diagnosis using task-related EEG signals.

To this end, the present study proposes a novel knowledge-guided framework (see Fig. 1), graph signal processing-graph convolutional networks (GSP-GCNs), to distinguish PD patients from healthy individuals using large-scale EEG network obtained from a specific task involving the manipulation of auditory feedback during vocal production. This framework consists of four sequential components: GSP, the graph-network module, the classifier, and the interpretable model. Firstly, the GSP module

analyzes and processes the large-scale EEG networks to identify dynamic connectivity patterns. Subsequently, the graph-network module captures these connectivity patterns as key features for classification. The classifier component then utilizes these extracted features to discriminate PD patients from healthy individuals. Lastly, the interpretable model is incorporated to enhance the interpretability of the framework by providing a global visualization of essential learned features of the model and aligning them with voice-related EEG microstates characteristics. By adopting this innovative approach, our GSP-GCNs framework aims to provide illustrative information for facilitating the use of the deep learning model in PD diagnosis with task-related EEG data.

RESULTS

Comparison of classified performance

The present study proposed four graph-network-based models: PCC+GCNs, PLV+GCNs, PCC + GSP-GCNs, and PLV + GSP-GCNs. The PCC + GSP-GCNs model constructs the brain network using PCC features and employs the GSP-GCNs model as the classifier. The PLV + GSP-GCNs model is similar to the PCC + GSP-GCNs model, except that it uses PLV features to construct brain networks. In contrast, the PCC+GCNs and PLV+GCNs models do not include the GSP module based on the aforementioned methods.

Table 1 provides the classification performance achieved by our proposed models and baseline models based on voice-related EEG signals, and Fig. 2 shows the results of non-parametric Mann-Whitney tests with Bonferroni correction that compare the classification performance (ACC and AUC) of these models in 5-fold cross validation. Our PCC + GSP-GCNs model outperformed

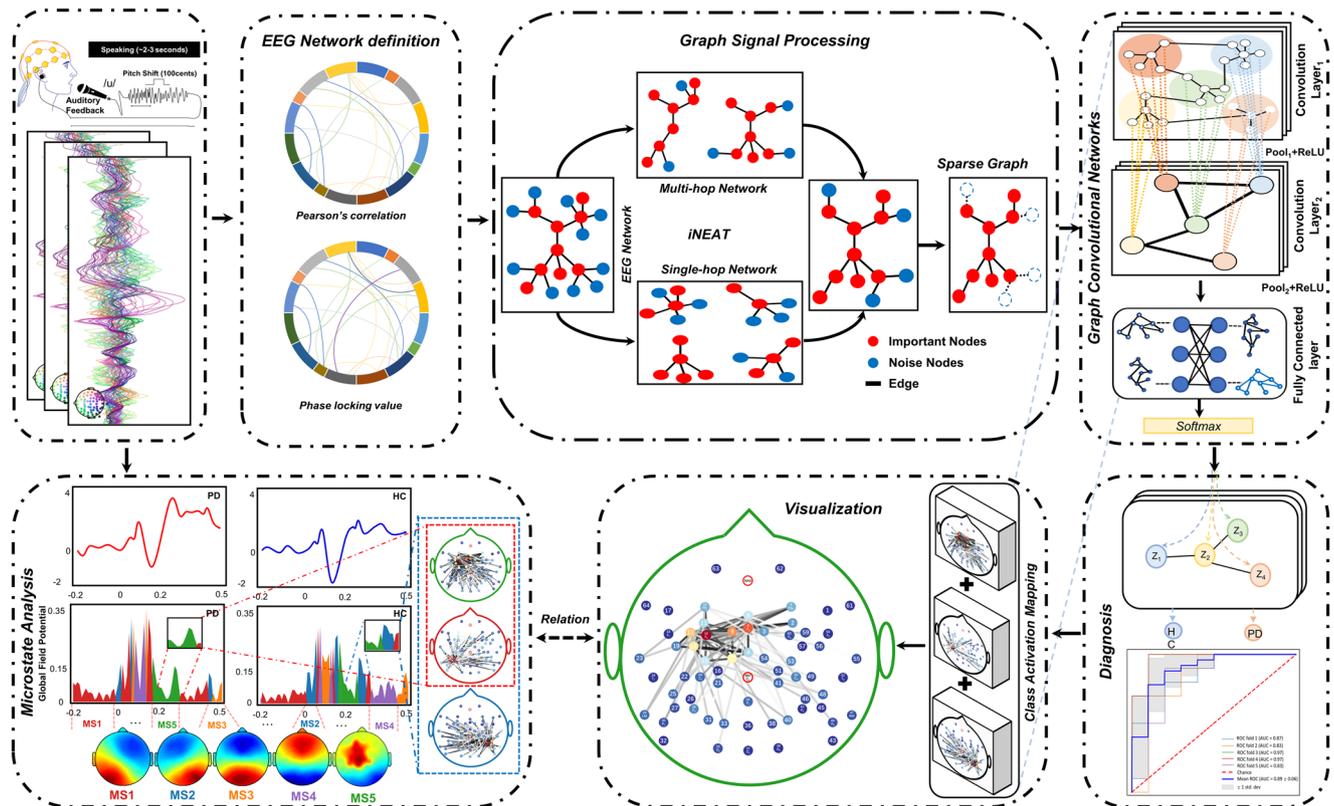


Fig. 1 The schematic diagram of the GSP-GCNs model designed for PD diagnosis. The GSP module is responsible for acquiring collaborative graph information of both single-hop and multi-hop networks from the EEG signals, while the GCNs module automatically learns the graph structure by aggregating the information of neighboring nodes. INA incomplete network alignment, GCN graph convolutional network, GSP graph signal processing.

Table 1. Classification performance and computational complexity of different deep learning models based on voice-related EEG signals, where k represents the convolution kernel size, d represents the time series length, and n represents the number of channels.

Models		Evaluation Criterion				
		ACC	AUC	Sensitivity	1-Specificity	Computational Complexity
Baseline Models	CNN	79.6%	78.1%	80.2%	77.5%	$4*O(k*n*d^2)+O(n^2)$
	RNN	75.2%	74.2%	80.1%	75.3%	$O(n*d^2)+O(n^2)$
	2D-CNN-RNN	81.6%	79.5%	78.6%	80.2%	$O(k*n*d^2)+O(n*d^2)+O(n^2)$
	3D-CNN-RNN	82.1%	81.2%	83.1%	78.2%	$2*O(k*n*d^2)+O(n*d^2)+O(n^2)$
	EEGNet	82.3%	81.2%	81.5%	79.6%	$3*O(k*n*d^2)+O(n^2)$
	Cascade model	80.4%	83.2%	82.1%	77.2%	$3*O(k*n*d^2)+2*O(n*d^2)+2*O(n^2)$
	Parallel model	81.1%	80.2%	82.3%	79.8%	$3*O(k*n*d^2)+O(n^2)$
GCNs	PCC+GCNs	84.1%	83.1%	78.2%	86.0%	$2*O(n\log n)+O(n^2)$
	PLV+GCNs	84.8%	82.2%	79.1%	84.3%	$2*O(n\log n)+O(n^2)$
	PLV + GSP-GCNs	88.9%	87.5%	86.1%	86.2%	$2*O(n\log n)+O(n^2)$
	PCC + GSP-GCNs	90.2%	89.1%	84.0%	88.4%	$2*O(n\log n)+O(n^2)$

ACC accuracy, AUC area under curve, CNN convolutional neural network, RNN recurrent neural networks, GSP graph signal processing, GCN graph convolutional network, PCC Pearson correlation coefficient, PLV phase locking value.

the EEGNet model significantly, with an accuracy increase of 7.9% (82.3% vs. 90.2%), an AUC increase of 7.9% (81.2% vs. 89.1%), a sensitivity boost of 2.5% (81.5% vs. 84.0%), and a specificity enhancement of 8.8% (79.6% vs. 88.4%). As well, our proposed graph-based models exhibit superior performance compared to previous methods that did not use graph learning²⁶. Specifically, the PLV + GSP-GCNs model improved those metrics by 8.2%, 7.7%, 3.6%, and 10.2%, compared to the 3D-CNN-RNN model. In addition, Table 1 shows the computational complexity of our proposed models and baseline models. The computational complexity of the GCNs model is $2*O(n\log n)+O(n^2)$ ^{32,33} while the computational complexity of the baseline models ranges from $O(n*d^2)+O(n^2)$ to $3*O(k*n*d^2)+2*O(n*d^2)+2*O(n^2)$ ^{26,34–36}, where k denotes the convolution kernel size (greater than 3), d denotes the time series length ($d = 700$), and n denotes the number of channels ($n = 64$). Accordingly, our GCNs model shows a considerably lower computational complexity than all baseline models.

Notably, Table 1 and Fig. 2 also show a comparison of classification performance across our proposed models. The GSP-GCNs models exhibited significantly superior performance compared to the GCNs models, regardless of the feature type. For example, the PCC + GSP-GCNs model achieved higher ACC, AUC, sensitivity, and specificity than the PCC+GCNs model by 6.1% (84.1% vs. 90.2%), 6.0% (83.1% vs. 89.1%), 5.8% (78.2% vs. 84%), and 2.4% (86.0% vs. 88.4%), respectively. Similarly, the PLV + GSP-GCNs model improved these metrics by 4.1% (84.8% vs. 88.9%), 5.3% (82.2% vs. 87.5%), 7.0% (79.1% vs. 86.1%), and 1.9% (84.3% vs. 86.2%) compared to the PLV+GCNs model. Furthermore, we employed a cross-validation approach to verify the out-of-distribution detection capability and stability of the GCNs model. Figure 3 indicates a consistently low variance of receiver operating characteristic (ROC) values, all remaining below 0.08 for our proposed models. These results highlight the potential of GSP processing in enhancing the classification performance by balancing the local and global information of single-hop and multi-hop networks through graph aggregation.

Interpretable GSP-GCN model

The interpretability of the GSP-GCNs framework was assessed in the present study, aiming to capture global frequency-spatial-temporal dependencies in the large-scale EEG network and extract essential information for decoding tasks from the time series data.

We used a modified CAM method to visualize the global representation learned by our models and generate their saliency maps in the context of vocal pitch regulation. Figure 4 illustrates the discriminative distributions of large-scale EEG networks obtained from the FAF task and learned by the PCC+GCNs and the PCC + GSP-GCNs models. The most significant discriminative distributions were located in the left ventral premotor cortex (vPMC), superior temporal gyrus (STG), and Broca's area. These results suggest that the GSP-GCNs models have the capability to capture the intrinsic representation of the major brain activity difference during vocal motor control between PD patients and healthy controls.

Moreover, we performed microstate analysis on the voice-related EEG signals from PD patients and healthy controls and identified five distinct microstates (MS1–MS5) with high energy fluctuation in the time frame of 0–300 ms (see Fig. 5). The most prominent microstate was the MS5, located in the electrodes near the left vPMC and Broca's area in the time frame of 205–315 ms. This temporal correspondence aligns with the large-scale EEG networks associated with the P2 component. The microstate transitioned from MS5 to MS2 during the 260–300 ms period of P2, with MS2 located near the right STG. The network identified in the MS5 microstate resembles the discriminative distributions learned by the GSP-GCNs model, further strengthening the interpretability of our proposed models.

DISCUSSION

The present study proposed a novel, interpretable deep learning framework based on GSP-GCNs to distinguish from PD patients from healthy controls using voice-related EEG signals. By incorporating both local and global information from single-hop and multi-hop networks, our GSP-GCNs model achieved an averaged classification accuracy of 90.2% and exhibited a performance improvement of 9.5% compared to other deep learning models. Moreover, the interpretability analysis of our GSP-GCNs models revealed discriminative distributions of large-scale EEG networks and topographic map of the microstate MS5 in the P2 time window, primarily located in the left vPMC, STG, and Broca's area that have been implicated in PD-related motor speech disorders. Overall, our proposed GSP-GCNs models offer a valuable tool for PD diagnosis based on the interpretable results derived from large-scale voice-related EEG networks.

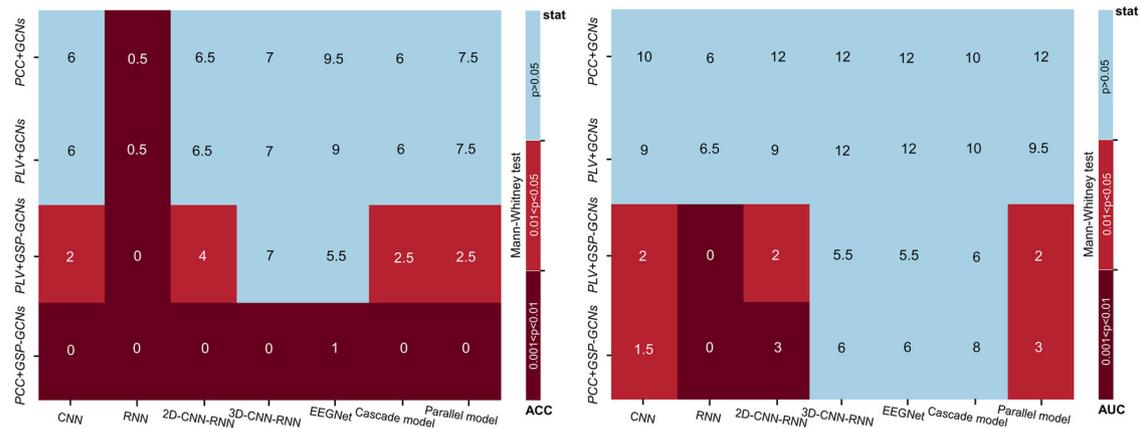


Fig. 2 Heatmaps showing statistical differences in classification performance between models. The left and right panels show the results of non-parametric Mann-Whitney tests comparing the GCNs model and conventional deep learning models in 5-fold cross validation for classification performance corresponding to ACC and AUC, respectively. Each cell in the heatmap represents the statistical significance of differences between the models, with light blue indicating no significance, red indicating a significant difference, and dark red denoting a higher level of significance.

A growing body of studies has concentrated on the development of PD diagnosis by integrating resting-state EEG with machine learning techniques^{12–14,37–39}. Recently, several deep learning models have been proposed for PD diagnosis using different resting-state EEG datasets (e.g. the public UNM dataset, public SanDiego dataset). For example, Oh et al.¹⁴ proposed a thirteen-layer CNN model while Lee et al.³⁹ proposed a hybrid CNN-RNN model, achieving classification accuracies of 88.25% and 99.2% in distinguishing PD patients from healthy controls, respectively. In addition, Shah et al.³⁸ developed a dynamical system generated hybrid network (DGHNet) for successful categorization of on-medication vs. off-medication PD patients with a classification accuracy of 99.2%. Nevertheless, the practicality of these models is constrained by the prolonged training period, the large datasets requirement, and the lack of interpretability of specific EEG features in PD diagnosis. Furthermore, the stochastic nature of resting-state EEG poses challenges in capturing the distinct patterns of PD-related brain activity.

In attempts to address these limitations, Shi et al.²⁶ developed four deep learning architectures (i.e. CNN, RNN, 2D-CNN-RNN, 3D-CNN-RNN) using voice-related EEG signals during vocal pitch regulation for PD diagnosis and achieved classification accuracies ranging from 76.00% to 82.89%. Hassin-Baer et al.⁴⁰ applied machine learning models for the diagnosis of early-stage PD using the event-related EEG signals during visual Go/No-Go and auditory Oddball cognitive tasks, achieving an AUC of 0.79 and identifying a total of 15 EEG features. In the present study, our proposed the GSP-GCNs models used voice-related EEG signals similar to Shi et al.²⁶ and achieved classification accuracies ranging from 84.1% to 90.2%, indicating a remarkable performance superiority over those models proposed by Shi et al.²⁶. Furthermore, our PCC + GSP-GCNs model also outperformed other models (see Table 1 and Fig. 2), including the EEGNet model, that do not use graph learning. Taken together, these findings demonstrate the effectiveness of deep learning models with graph learning and voice-related EEG signals for PD diagnosis.

Our proposed GSP-GCNs models, integrated with voice-related EEG, present a significant advancement over other deep learning models for PD diagnosis in several aspects. A significant strength lies in their capacity to combine the global and local properties of brain functional networks using single-hop and higher-order networks, enabling the aggregation of information from the nodes of the functional brain networks and the extraction of features for graph learning. Another strength is the use of GCN models for graph data, which have advantages over other convolutional

neural networks, such as capturing complex node relationships, adapting to different graph properties, and handling large-scale network data. In contrast, the CNN-based approaches only focus on local features due to their limited perceptual field but ignore the global correlations among EEG signals. Similarly, the CNN-RNN approaches only capture the sequential relationships of EEG signals while overlooking the relevance of the EEG connectivity network. On the other hand, resting-state EEG signals are stochastic in nature and fail to capture the functional specificity of brain activity related to PD when compared to task-related EEG signals. The present study used a task that involves auditory-motor integration for vocal pitch regulation, which reveals the neural processes involved in the coordination of sensory and motor systems for speech production. These processes have been demonstrated to be impaired in PD^{21–25}. Therefore, our GSP-GCNs models provide a novel way for PD diagnosis using voice-related EEG signals and deep learning techniques, which is more reliable and feasible than previous methods.

More importantly, the analysis that revealed comprehensive global frequency-spatial-temporal dependencies within the extensive EEG network demonstrated the interpretability of our proposed models, providing insights into how our models learn from voice-related EEG signals and how they relate to the neural processes underlying impaired auditory-vocal integration in PD. Upon visualizing the overall representation acquired by our models and their corresponding saliency maps during vocal pitch regulation, the most prominent discriminatory distributions were located in the left vPMC, STG, and Broca's area. These regions have been well-established as integral to speech motor control^{41–46}. This consistency suggests that our GSP-GCNs models capture distinct patterns of brain activity during vocal motor control between PD patients and healthy controls.

EEG microstate analysis has been proven valuable in the diagnosis of Alzheimer's disease⁴⁷ and schizophrenia¹³ by revealing the spatiotemporal dynamics of brain activity. In the present study, the microstate analysis of voice-related EEG signals identified five distinct microstates (MS1–MS5), with MS5 being of particular significance. Our observation of distinct MS5 microstate between PD patients and healthy controls is in line with previous studies that have shown associations of specific EEG microstates with vocal tract muscles and motor cortex activity^{48,49} as well as speech fluency in PD patients⁵⁰. MS5 corresponds to the large-scale EEG networks that align with the P2 component (205–315 ms). This ERP component has been thought to reflect a complex stage of auditory-motor transformation for controlling

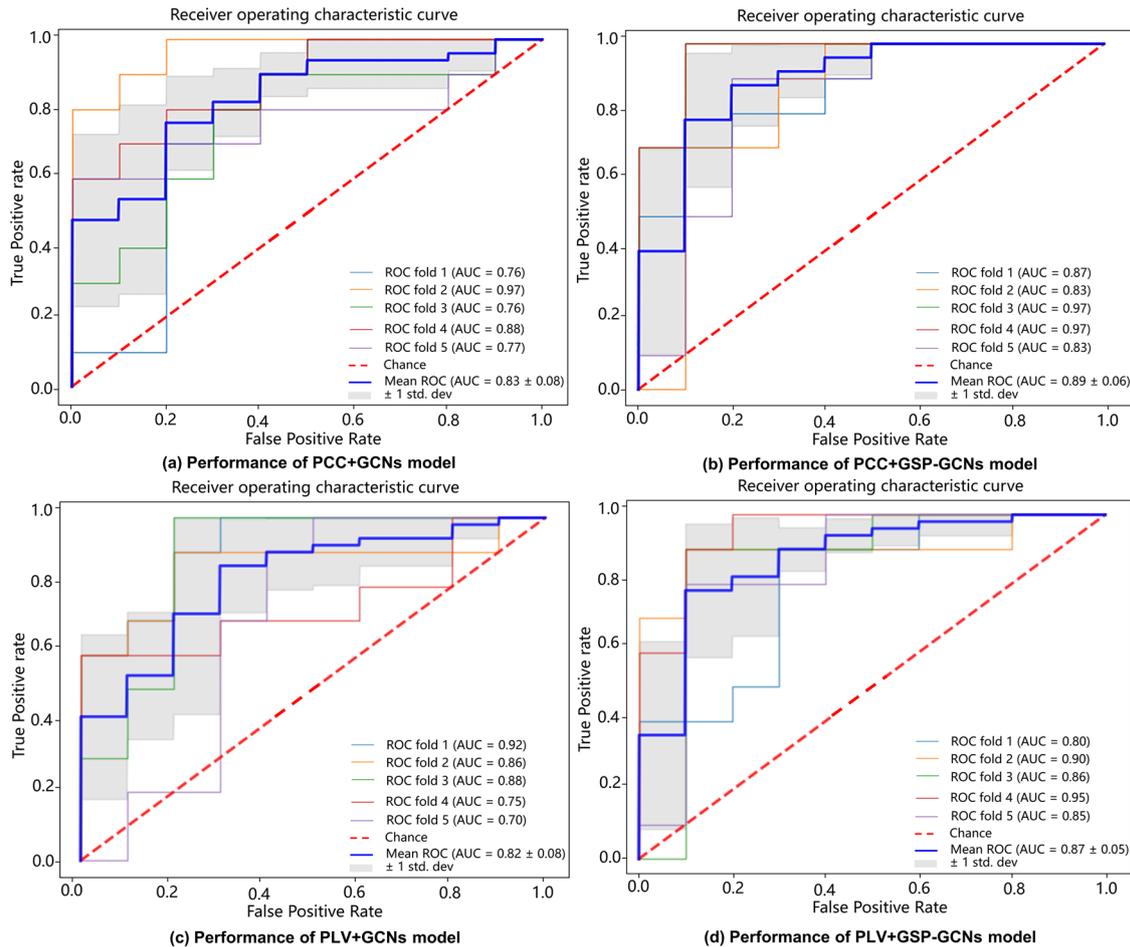


Fig. 3 The ROC curves of the GCNs models. Panels a–d show the ROC curves for the PCC+GCNs, PLV+GCNs, PCC + GSP-GCNs, and PLV + GSP-GCNs models, respectively. The solid blue line represents the average performance across different test sets using 5-fold cross-validation, while the other colored lines show the performance of individual test set. The gray shaded area indicates the variance range of the average ROC curve.

vocal production that demands higher-level cognitive processing^{51,52}. In particular, enhanced P2 responses to vocal pitch perturbations have been linked to impaired auditory-vocal integration in PD patients when compared to healthy controls^{23–25}. Notably, the spatial localization of MS5 coincided with the left vPMC and Broca’s area (see Fig. 4), regions that have been recognized as significant contributors to enhanced P2 responses during vocal pitch regulation in PD patients compared to healthy controls²³. More interestingly, this network representation of MS5 microstate also showed a resemblance to the discriminative distributions learned by the GSP-GCNs model, lending further support to the interpretability of our proposed models to unravel specific brain patterns during speech tasks relevant to PD. Therefore, this interpretability analysis highlights the novelty and significance of our approach in facilitating a more profound understanding of the neural mechanisms underlying PD. Such interpretability is crucial^{27,28}, ensuring that the models are not just “black boxes” but provide meaningful insights into the neural dynamics underpinning PD.

The proposed GSP-GCNs models have important clinical implications for PD diagnosis. They offer the potential to improve the diagnostic accuracy by extracting interpretable features from large-scale voice-related EEG networks. They also contribute to reducing the subjective bias and variability across patients, thereby promoting a more objective and consistent assessment of PD. Moreover, they hold the promise to facilitate the treatment

of speech impairment in PD by modulating activity of the brain regions observed in microstate analysis through the use of non-invasive brain stimulation techniques such as transcranial magnetic or electrical stimulation^{25,53–55}. Therefore, our GSP-GCNs models not only represent an advancement in PD diagnosis but also may pave the way for effective treatment approaches.

Online detection of clinical diseases is a challenging task that requires high accuracy and low latency for brain signals. However, most existing EEG systems for PD diagnosis are based on offline analysis of resting-state data, which suffers from high noise and low stability in real-time EEG signals due to factors such as subject concentration and device inconsistencies. Previous studies have demonstrated the effectiveness and adaptability of GCN models in real-world scenarios across various domains, such as online recommender systems⁵⁶, traffic flow prediction⁵⁷, and online animal tracking⁵⁸. The present study proposed a hybrid approach that combined the computing capabilities of offline systems with the real-time responsiveness of online systems based on the voice-related EEG signals. In this framework, the offline system constructed the large-scale EEG networks to extract the features of specific ERP components, while the online system applied the GSP-GCNs model to enable real-time evaluation. This strategy achieved a balance between high accuracy and low latency in classification, allowing us to detect unique patterns within task-related EEG signals for online PD diagnosis.

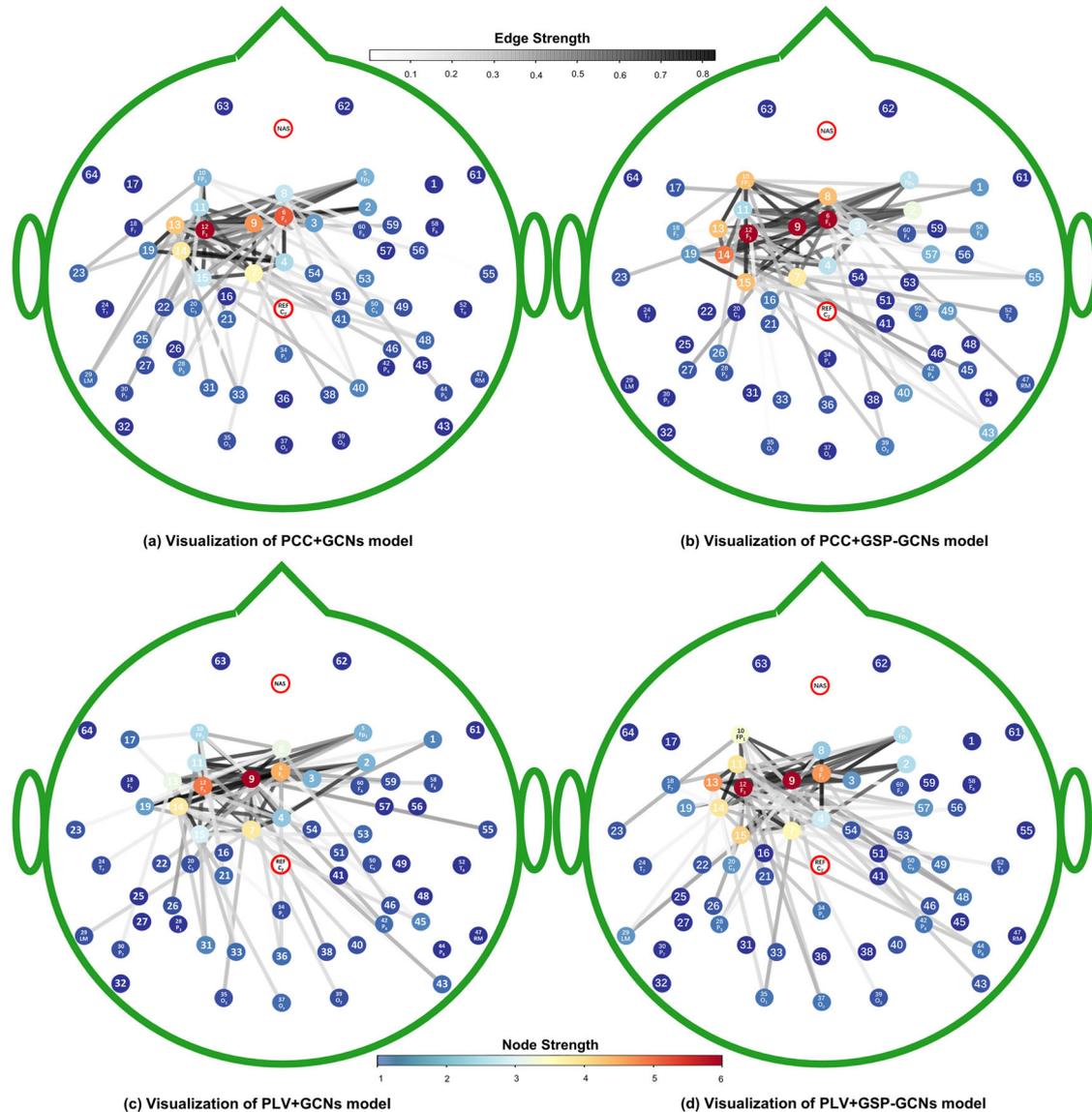


Fig. 4 The saliency maps of the discriminative biomarkers learned by the GCNs models for the FAF task. The numbers ranging from 1 to 64 correspond to the electrodes according to the international EEG 10-20 system. Panels a–d show the distribution of the contribution network for the PCC+GCNs model, the PCC + GSP-GCNs model, the PLV+GCNs model, and the PLV + GSP-GCNs model, respectively. Nodes with a redder hue indicate greater contributions to the model, and edges with a blacker hue indicate higher contributions.

This claim is supported by a posteriori verification of the proposed models (see Supplementary Fig. 1), which compares the performance of different deep learning models based on the random selection of trials from a pool of 100 trials for a voting selection process. Remarkably, the post-validation performance of our models exhibits only slight declines (4.1–4.9%) when compared to the results presented in Table 1. In contrast, previous studies^{59,60} have reported that online detection systems based on the brain signals typically experience an approximate 10% reduction in accuracy compared to their offline counterparts. Moreover, the GSP-CNSs models outperform both conventional deep learning models and previous EEG-based online detection systems^{59,60} in stochastic posterior performance and computational complexity (see Table 1 and Fig. 2). Therefore, our proposed models have high potential for online detection scenarios due to their high accuracy, low computational complexity, and objective evaluation.

There are several limitations in the present study that warrant further investigation. First, our sample size was relatively small due

to difficulties in obtaining task-related EEG data under controlled experimental conditions and specialized equipment requirements, which may limit the generalizability and robustness of our models. Second, our models performed the connectivity analysis of large-scale EEG networks at the electrode level, which cannot provide precise anatomical sources that generate the EEG signals. Incorporating source localization analysis for constructing EEG connectivity network is therefore needed in future studies to provide a more accurate representation of functional interactions within the brain. Also, the exclusive use of GCN models in the present study may not be suitable for all types of graph data. Future studies should consider alternative methodologies (e.g. graph attention networks, graph neural ordinary differential equations) for graph learning. Lastly, our models focused on a specific task that involves vocal motor control for PD diagnosis. Other tasks, such as cognitive control, may reveal different brain activity patterns and network dynamics that can be informative for PD diagnosis.

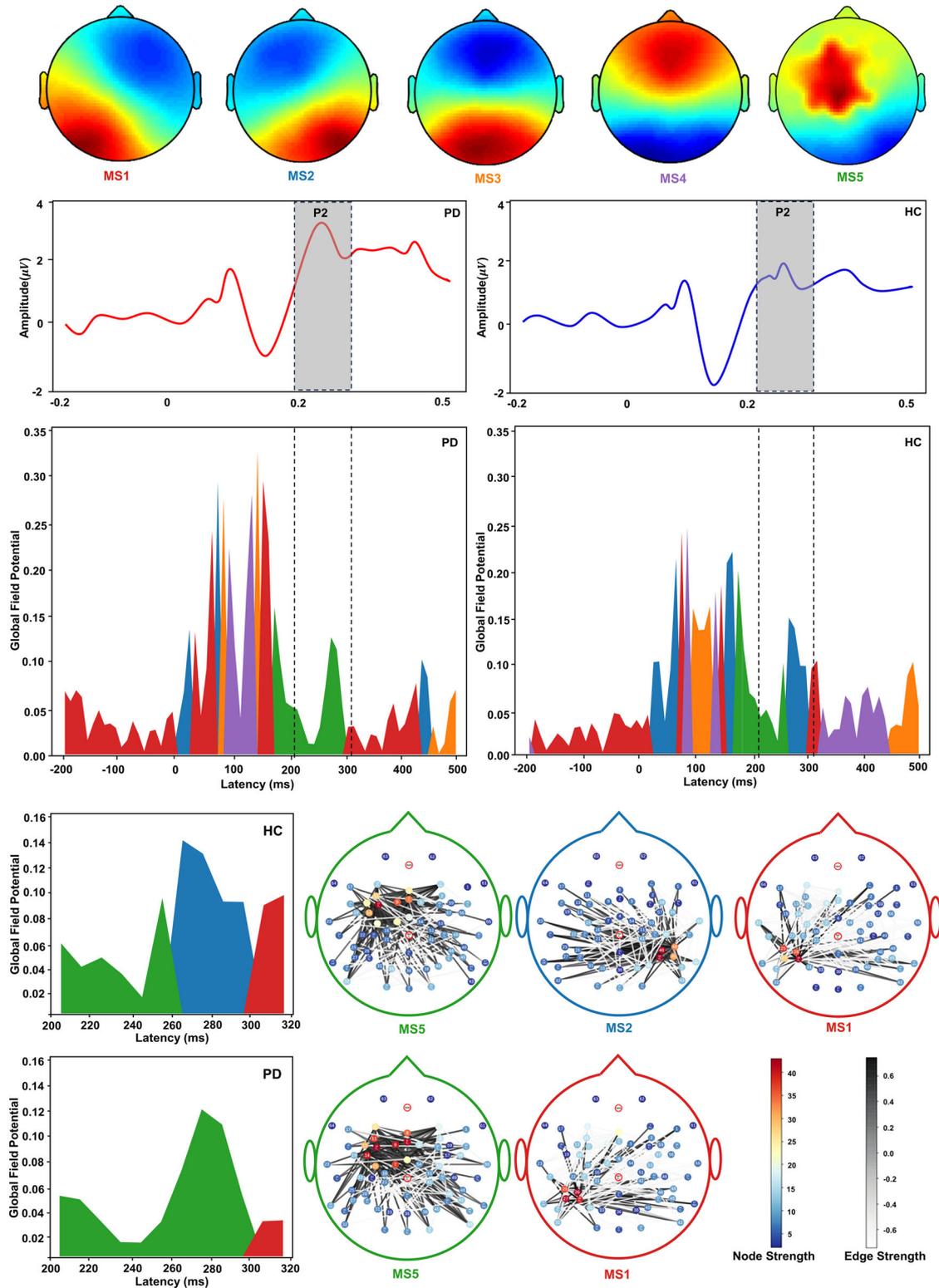


Fig. 5 The distribution of microstate classes corresponding to grand-averaged ERP in the FAF task for PD and HC groups. Different colors represent different microstates. Five microstates clustered by k-means algorithm are shown in the first row, average time potential fluctuations with emphasis on P2 component are shown in the second row, global field power is shown in the third row, and functional activity distributions within microstates and EEG networks associated with P2 component are shown in the bottom. GFP global field power, MS microstate.

In conclusion, the present study proposed a novel deep learning framework based on GSP-GCNs for PD diagnosis using voice-related EEG signals. Our models can capture the global frequency-spatial-temporal dependencies among large-scale EEG networks, achieving a remarkable 90.2% classification accuracy and outperforming other deep learning models by 9.5%. Moreover, our models revealed the discriminative distributions of large-scale EEG networks and topography of microstate MS5 for PD diagnosis in terms of interpretability. These findings highlight the promise of interpretable deep learning models with task-related EEG signals in advancing PD diagnosis.

METHODS

Task-related EEG Dataset

Fifty-two patients diagnosed with idiopathic PD (24 females and 28 males; mean age = 64.23 ± 5.30 years) according to the diagnostic criteria of the UK Parkinson's disease Society Brain Bank⁶¹ and forty-eight sex- and age-matched healthy controls (HC) (23 females and 25 males; mean age = 63.37 ± 5.41 years) participated in this study. All of them were right-handed, native Mandarin speakers. PD patients in the present study met the following inclusion criteria: no more than mild dementia (Mini-Mental State Examination [MMSE] > 26), no other neurological disease, no history of neurosurgical treatment, laryngeal surgery or pathology, swallowing disorders. PD patients were kept on their antiparkinsonian medication, but they were tested during their off-medication state (i.e. 12 h off anti-PD medication). All participants provided informed consent, and the research protocol was approved by the Institutional Review Board of The First Affiliated Hospital at Sun Yat-sen University in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Task-related EEG data were acquired during a task based on the frequency altered feedback (FAF) paradigm⁶². In brief, participants were instructed to produce a sustained vowel sound (/u/) for a duration of 5–6 s while hearing their voice unexpectedly pitch-shifted downwards by 200 cents (100 cents = one semitone) for a duration of 200 ms. Each vocalization consisted of 4–5 perturbations that were presented in a pseudorandomized manner. Participants produced 20–25 consecutive vocalizations, resulting in a total of 100 trials. More details regarding experimental designs for the FAF tasks can be found in previous studies^{23,25}.

While participants performed the vocal production experiment, the EEG signals were scalp-recorded using a 64-electrode Geodesic Sensor Net connected to a Net Amps 300 amplifier (Electrical Geodesics Inc.) at a sampling frequency of 1 kHz using NetStation software (v.4.5, Electrical Geodesics Inc.). During the offline analysis, the EEG signals were band-filtered (1–20 Hz) and segmented with a window of –200 ms before and 500 ms after the perturbation onset. An artifact detection procedure was applied to the segmented epochs to exclude those bad trials from further analysis. And artifact-free trials were re-referenced to the average of the electrodes on each mastoid, averaged, and baseline-corrected to generate an overall cortical ERP response.

System framework

Figure 1 shows the overall framework of knowledge-guide graph convolutional networks, which consists of four modules: EEG network, GSP, GCN, and interpretable module. Firstly, the EEG network module constructs a graph of dynamic brain activity using electrodes as nodes and functional connections between electrodes as edges^{63,64}. Subsequently, the GSP module applies a strategy based on incomplete network alignment (iNEAT)⁶⁵ and Sparse Graph⁶⁶ to reorganize the local and global information in the EEG network. Next, the GCNs module learns graph representations for personalized diagnosis by aggregating information from neighboring nodes, capturing intrinsic features from the complex EEG network for PD

diagnosis. In addition, the interpretable module uses a modified saliency map derived from the backpropagation algorithm⁶⁷ to visualize the prominent EEG network in each individual pattern and then compares it with the large-scale EEG network obtained from microstate analysis. The details of each module are described below.

EEG network definition. The EEG network is modeled by a graph $\mathbf{G}(V, E)$, where v_i represents the i th channel of EEG in a node set V while e_{ij} represents the strength of functional connection between nodes v_i and v_j in an edge set E . Specifically, a single-channel EEG data records a sequence of time-series $\mathbf{r}_{i \in [1, n]} = [\mathbf{s}_1^i, \mathbf{s}_2^i, \dots, \mathbf{s}_k^i] \in \mathbf{R}^k$, where $k = T \times f$ denotes the number of time points (f represents the frequency range, and T represents the consecutive order of time series). The EEG signal can be represented as a tensor $\mathbf{X}_T = [\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n] \in \mathbf{R}^{n \times k}$, and $\mathbf{G}(V, E)$ can be represented as an adjacency matrix. Functional connectivity matrices (V) are calculated using Pearson Correlation Coefficient (PCC) or Phase Locking Value (PLV). The PCC between brain signals in different channels is defined by Eq. (1), where $\mathbf{r}_i^{f_p}(\mathbf{t})$ corresponds to the low energy fluctuation within the frequency band f_p (4–12 Hz) from the i th electrode of the EEG signals.

$$PCC(\mathbf{r}_i^{f_p}, \mathbf{r}_j^{f_p}) = \frac{1}{N_s} \sum_{k=1}^{N_s} \mathbf{r}_i^{f_p}(\mathbf{k}) \mathbf{r}_j^{f_p}(\mathbf{k}) \quad (1)$$

PLV, defined in Eq. (2), measures the synchronization between phases of brain regions.

$$PLV(\mathbf{r}_i^{f_p}, \mathbf{r}_j^{f_p}) = \left| \left\langle e^{im\Delta\varnothing_r(\mathbf{t})} \right\rangle \right| = \left| \frac{1}{N_s} \sum_{k=1}^{N_s} e^{im(\varnothing_i(\mathbf{k}) - \varnothing_j(\mathbf{k}))} \right| \quad (2)$$

$\Delta\varnothing_r(\mathbf{t})$ is the instantaneous phase calculated by Hilbert transform from the original signal $\mathbf{r}^f(\mathbf{t})$.

The traditional GCNs focus on the local network structure rather than the global network distribution. To overcome this limitation, the present study incorporated a GSP method, including the iNEAT algorithm and Sparse Graph, into the GCNs. This integration allows us to capture both the local and global information within the EEG network. The iNEAT component selects edge features based on the hub and link properties of nodes, while the sparse graph component adjusts node weights according to edge weights. These methods enable a reorganization of the EEG network to obtain a new network that contains important nodes and edges while reducing or eliminating noisy ones. This reorganization is an important step in the GSP process for enhancing the data quality, as evidenced in previous studies on graph learning^{68,69}. Although the iNEAT algorithm integrates both the local and global information within the large-scale EEG network, it may result in an issue of over completeness in fusion information. To overcome this limitation, the present study introduced the Sparse Graph operation, which performs sparse decomposition and dictionary generation on the graph signal, to remove redundant information.

Two graph networks, \mathbf{G}_1 from a single-hop network ($k = 1$) and \mathbf{G}_2 from a multi-hop network ($k = 8$), were generated for each individual using the k -nearest neighbor method⁶⁵. The adjacency matrices \mathbf{A}_1 and \mathbf{A}_2 were calculated to obtain the permutation matrix \mathbf{S} . To optimize the global graph information of the single-hop network, a graph matching-based method was proposed to solve a Non-deterministic Polynomial problem according to the principle of topology consistency^{70,71}.

$$\min_S \|\mathbf{A}_2 - \mathbf{S}^T \mathbf{A}_1 \mathbf{S}\|_F^2 \quad (3)$$

where $\|\cdot\|_F$ is the Frobenius norm of the corresponding matrix.

The random walk-based method was used to capture

Table 2. Number of training parameters of the GCN model according to the Chebyshev polynomial order of each layer.

Layers	Shape of weight tensor	Shape of bias	Number of parameters
Convolution layer ₁	[$\mathbf{K}_1, 64, 64$]	[64]	$4096 \times \mathbf{K}_1 + 64$
Pool ₁	[1×2]	[1×2]	4
Convolution layer ₂	[$\mathbf{K}_2, 64, 32$]	[32]	$2048 \times \mathbf{K}_2 + 32$
Pool ₂	[2×1]	[2×1]	4
Fully Connected layer	[32, 32]	[32]	1056

\mathbf{K}_1 and \mathbf{K}_2 are the coefficients of the Chebyshev polynomial expansion.

collaborative graph information from both single- and multi-hop networks based on the Kronecker product graph^{72,73}.

$$\mathbf{S} = \alpha(\mathbf{A}_1 \otimes \mathbf{A}_2)\mathbf{S} + (1 - \alpha)\mathbf{h} \quad (4)$$

where \mathbf{h} is the vectorization of the prior similarity matrix \mathbf{H} via the sine function between mean adjacency matrices in different groups, and \otimes is the operation of Kronecker product graph.

The iNEAT algorithm, which combines the strengths of the graph matching-based method and the random walk-based method⁶⁵, was used in the present study. This algorithm effectively integrates the graph structure information from both single- and multi-hop networks within large-scale EEG networks. The optimization objective function is defined as follows:

$$\min_{\mathbf{s}} \alpha \mathbf{s}^T (\mathbf{D} - \mathbf{A}_1 \otimes \mathbf{A}_2) \mathbf{s} + (1 - \alpha) \|\mathbf{D}^{\frac{1}{2}}(\mathbf{s} - \mathbf{h})\|_F^2 \quad (5)$$

where \mathbf{s} represents the vectorization of the similarity matrix \mathbf{S} . $\mathbf{D} = \mathbf{D}_1 \otimes \mathbf{D}_2$ and \mathbf{D}_1 and \mathbf{D}_2 are the diagonal degree matrix corresponding to \mathbf{A}_1 , and \mathbf{A}_2 . In addition, a permutation matrix was used to reorganize the EEG signals through channel-wise operations at the individual level.

The Sparse Graph operation, defined in Eq. (6), aims to enhance the discriminative ability between groups and reduce the standard deviation across the trials. In Eq. (6), the parameter of \mathbf{x}_i^e represents the EEG signal of i th electrode from the i th subject, while the graph matrix \mathbf{w}_i^e is calculated through the iNEAT algorithm to learn the collaborative graph information.

$$\begin{aligned} \mathbf{W}^e = \arg \min_{\mathbf{W}^e} & \sum_{i=1}^n \left(\frac{1}{2} \|\mathbf{x}_i^e - \mathbf{X}_i^e \mathbf{w}_i^e\|_2^2 \right) + \lambda_1 \|\mathbf{B}_g^e \odot \mathbf{W}^e\|_{2,1} + \lambda_2 \sum_{i,j=1}^n s_{ij}^e \|\mathbf{w}_{gi}^e - \mathbf{w}_{gj}^e\|_2^2 \\ & + \lambda_3 \|\mathbf{B}_t^e \odot \mathbf{W}^e\|_{2,1} + \lambda_4 \sum_{i,j=1}^n s_{ij}^e \|\mathbf{w}_{ti}^e - \mathbf{w}_{tj}^e\|_2^2 \end{aligned} \quad (6)$$

Where $\mathbf{B}_g^e = [\mathbf{b}_{g1}^e, \dots, \mathbf{b}_{gi}^e, \dots, \mathbf{b}_{gn}^e]$ is a weighting matrix with elements being $\mathbf{b}_{gi}^e = [b_{gi}^{e,1}, \dots, b_{gi}^{e,e-1}, b_{gi}^{e,e}, b_{gi}^{e,e+1}, \dots, b_{gi}^{e,E}]$. Similarly, the matrix \mathbf{B}_t^e is defined. s_{ij}^e and s_{ij}^e denote the similarity between i th and j th subject from different groups.

Graph Convolutional Networks. In contrast to the spectrum-based approach based on the CNNs, the GCNs use the graph structure to intelligently aggregate information from neighboring nodes. Spectral analysis of graph signals decomposes EEG signals into multi-frequency graph modes and identifies the distribution of EEG signals using Laplacian maps from the spatial domain to the spectral domain. Laplacian operator, denoted as \mathbf{L} , is defined as follows:

$$\mathbf{L} = \mathbf{D} - \mathbf{A} \in \mathbb{R}^{N \times N} \quad (7)$$

$$\mathbf{L} = \mathbf{I}_N - \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}}, \mathbf{D} \in \mathbb{R}^{N \times N} \quad (8)$$

where \mathbf{D} is the diagonal degree matrix, and \mathbf{I}_N is an identity matrix of size N .

The graph signal can be defined through the Fourier transform⁷⁴.

$$\chi = \mathbf{U}^T \mathbf{x}, \mathbf{x} = \mathbf{U} \chi \quad (9)$$

where $\mathbf{U} = [\mathbf{u}_0, \dots, \mathbf{u}_{N-1}] \in \mathbb{R}^{N \times N}$ is calculated through the eigenvector decomposition of \mathbf{L} .

In contrast to traditional methods that calculate a weighted sum of spatial neighbors in the Euclidean space, the ChebNet employed in the present study applied graph filters to generate a linear combination of graph Fourier modes across different frequencies. The convolution operation between two graph signals \mathbf{x} and \mathbf{y} can be expressed through the graph \ast_g :

$$\mathbf{x} \ast_g \mathbf{y} = \mathbf{U} \left((\mathbf{U}^T \mathbf{x}) \odot (\mathbf{U}^T \mathbf{y}) \right) \quad (10)$$

where \odot represents the element-wise Hadamard product.

Since $\mathbf{L} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$ and $\mathbf{\Lambda} = \text{diag}([\lambda_0, \dots, \lambda_{N-1}])$, we defined g_θ as the filter with parameter θ . The filtering process of signal \mathbf{x} can be expressed as:

$$\mathbf{y} = g_\theta(\mathbf{L})\mathbf{x} = g_\theta(\mathbf{U} \mathbf{\Lambda} \mathbf{U}^T)\mathbf{x} = \mathbf{U} g_\theta(\mathbf{\Lambda}) \mathbf{U}^T \mathbf{x} \quad (11)$$

To avoid calculating the spectral decomposition of the graph Laplacian, the ChebNet model used a truncated expansion of the Chebychev polynomials⁷⁴.

$$g_\theta(\mathbf{\Lambda}) = \sum_{k=0}^{K-1} \theta_k T_k(\mathbf{\Lambda}) \quad (12)$$

$$\mathbf{y} = \mathbf{U} g_\theta(\mathbf{\Lambda}) \mathbf{U}^T \mathbf{x} = \sum_{k=0}^{K-1} \theta_k \mathbf{U} T_k(\mathbf{\Lambda}) \mathbf{U}^T \mathbf{x} = \sum_{k=0}^{K-1} \theta_k T_k(\mathbf{L}^\sim) \mathbf{x} \quad (13)$$

where \mathbf{L}^\sim is scaled Laplacian: $\mathbf{L}^\sim = 2\mathbf{L}/\lambda_{\max} - \mathbf{I}_N$.

Interpretable GSP-GCNs model. The present study used two methods to interpret the essential features derived from the GSP-GCNs model, including the examination of the features learned by the GSP-GCNs model and an analysis of the features of the large-scale EEG network through microstate analysis. The GSP-GCNs model was designed to capture global temporal dependencies within EEG data, enabling the identification of crucial information for decoding tasks from time series. While topography and Gradient-weighted Class Activation Mapping (CAM)⁷⁵ have been used to reveal the global representation learned by deep learning models in motor imagery dataset¹², the visualization of the gradients using the CAM often suffers from high levels of noise. To address this issue, the present study used a deconvolution approach that suppresses the flow of gradients through neurons³⁰. Specifically, for a given layer l in the graph signal χ^l and its gradient R^l , the overwritten gradient $\nabla \chi^l R^l$ can be calculated as follows:

$$\nabla \chi^l R^l = (\chi^l > 0) \odot (\nabla \chi^{l+1} R^{l+1} > 0) \odot \nabla \chi^{l+1} R^{l+1} \quad (14)$$

To generate the saliency map, the present study started from the output layer of a pre-trained model and propagated the gradients at each layer using the chain rule until reaching the input layer. This process allows us to visualize the salient regions that significantly contribute to the model's predictions, providing valuable insights into the interpreted features.

The microstate analysis of ERPs aims to investigate whether the GSP-GCNs model can capture distinct neural representations between PD patients and healthy controls when they produce vocal compensations for pitch feedback perturbations. Microstate analysis is a method that assesses the functional dynamics of large-scale brain networks by identifying the stable topographic patterns of the EEG signals^{47,50}. This analysis was performed using the MNE-python toolbox (<https://mne.tools/stable/index.html>), which involves calculating the global field power (GFP) of each

participant's ERP followed by clustering the topographic map underlying the GFP using the k-means algorithm from the *sklearn* toolbox. The Krzanowski-Lai criterion^{76,77} was used to determine the optimal number of microstate classes due to its suitability for selecting topographic map classifications based on quality indicators and global explained variance.

Experimental design

The GSP-GCNs model was implemented using the Pytorch toolkit with a 5-fold cross-validation strategy (https://github.com/ShuzhiZhao/ERP_GCN). The model parameters were optimized using the Adam optimizer with gradient descent and the cross-entropy loss function. The network had three GCN layers (two hidden layers and one fully connected layer) and a learning rate of the network was 10^{-5} . Table 2 shows the training parameters of the GCN model according to the Chebyshev polynomial order of each layer. A dropout rate of 0.35 was applied to prevent overfitting. The final output of the GSP-GCNs model was an M -dimensional vector obtained through the *Softmax* function, where M represents the number of EEG categories. The cross-entropy loss function, as defined in Eq. (15), was used to evaluate the model performance, where y and y^{\wedge} represent the ground truth and predicted label, respectively. N_b denotes the number of trials in a batch.

$$\mathcal{L} = -\frac{1}{N_b} \sum_{i=1}^{N_b} \sum_{c=1}^M y \log(y^{\wedge}) \quad (15)$$

To evaluate our proposed method, we compared it with several state-of-the-art approaches and performed ablation studies to show the impact of GSP. These baseline models used for comparison were Hybrid Convolutional Recurrent Neural Networks (CNN, RNN, 2D CNN-RNN, and 3D-CNN-RNN)²⁶, EEGNet³⁴, and CRNN (Cascade and Parallel model)³⁵. The performance was evaluated using ACC, AUC, sensitivity, and specificity (1-specificity).

Posteriori verification

A posteriori verification was performed to compare the classification performance of different deep learning models based on the random selection of trials from a pool of 100 trials for a voting selection process. The procedure consists the following steps:

Step 1: For each subject, 10 trials' classification labels are randomly selected from a set of 100 trials' classification labels $\mathbf{Y} = \{y_1, y_2, \dots, y_{10}\}$, where $y_i \in \{0, 1\}$ denotes the label of the i -th trial (0 represents PD patients, 1 represents healthy controls).

Step 2: A hard-voting strategy is used, where the total number of 0 labels \mathbf{N}_0 and the total number of 1 labels \mathbf{N}_1 in each subject's 10 trials are computed. If $\mathbf{N}_0 > \mathbf{N}_1$, the subject is classified as having PD and the procedure progresses to Step 4. If $\mathbf{N}_0 < \mathbf{N}_1$, the subject is classified as healthy, bypassing to Step 4. If $\mathbf{N}_0 = \mathbf{N}_1$, the procedure advances to Step 3.

Step 3: The Boyer-Moore algorithm is applied for majority voting. For example, if $\mathbf{Y} = \{0, 1, 1, 0, 1, 0, 0, 1, 1, 0\}$, thus $\mathbf{N}_0 = \mathbf{N}_1$ occurs. The Boyer-Moore algorithm's criterion is: the first trial vote is 0, the second trial vote is 1, the two votes are different and cancel each other; the second trial vote is 1, the third trial vote is 1, the two votes are the same, and the vote of \mathbf{BM}_1 increases by 1. We get $\mathbf{BM}_0 = 1$ and $\mathbf{BM}_1 = 2$, and classify the subject as healthy.

Step 4: Steps 1–3 are reiterated 100 times, and the mean and variance of the outcomes from these repetitions are calculated to provide a statistical overview of the classification performance.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

Anonymized data may be shared on request to the corresponding author for non-commercial use, subject to restrictions according to participant consent and data protection legislation.

CODE AVAILABILITY

Code for all data analysis is available from the corresponding author upon request and establishment of data sharing agreement between institutions.

Received: 23 August 2023; Accepted: 29 November 2023;

Published online: 05 January 2024

REFERENCES

- Selvaraj, S. & Piramanayagam, S. Impact of gene mutation in the development of Parkinson's disease. *Genes Dis.* **6**, 120–128 (2019).
- Armstrong, M. J. & Okun, M. S. Diagnosis and treatment of parkinson disease: a review. *JAMA* **323**, 548–560 (2020).
- Blauwendraat, C., Nalls, M. A. & Singleton, A. B. The genetic architecture of Parkinson's disease. *Lancet Neurol.* **19**, 170–178 (2020).
- Bloem, B. R., Okun, M. S. & Klein, C. Parkinson's disease. *Lancet* **397**, 2284–2303 (2021).
- Drotar, P. et al. Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease. *Artif. Intell. Med.* **67**, 39–46 (2016).
- Miller, D. B. & O'Callaghan, J. P. Biomarkers of Parkinson's disease: present and future. *Metabolism.* **64**, S40–S46 (2015).
- Horak, F. B. & Mancini, M. Objective biomarkers of balance and gait for Parkinson's disease using body-worn sensors. *Mov. Disord.* **28**, 1544–1551 (2013).
- Wroge, T. J. et al. Parkinson's disease diagnosis using machine learning and voice. In: *2018 IEEE signal processing in medicine and biology symposium (SPMB)*. IEEE (2018).
- Chu, C. et al. Spatiotemporal EEG microstate analysis in drug-free patients with Parkinson's disease. *NeuroImage. Clin.* **25**, 102132 (2020).
- Yi, G. S., Wang, J., Deng, B. & Wei, X. L. Complexity of resting-state EEG activity in the patients with early-stage Parkinson's disease. *Cogn. Neurodyn.* **11**, 147–160 (2017).
- Moazami-Goudarzi, M., Sarnthein, J., Michels, L., Moukhtieva, R. & Jeanmonod, D. Enhanced frontal low and high frequency power and synchronization in the resting EEG of parkinsonian patients. *NeuroImage* **41**, 985–997 (2008).
- Song, Y., Zheng, Q., Liu, B. & Gao, X. EEG conformer: Convolutional transformer for EEG decoding and visualization. *IEEE Trans. Neural Syst. Rehabil. Eng.* (2022).
- Lillo, E., Mora, M. & Lucero, B. Automated diagnosis of schizophrenia using EEG microstates and deep convolutional neural network. *Expert. Syst. Appl.* **209**, 118236 (2022).
- Oh, S. L. et al. A deep learning approach for Parkinson's disease diagnosis from EEG signals. *Neural. Comput. Appl.* **32**, 10927–10933 (2020).
- Chaturvedi, M. et al. Quantitative EEG (QEEG) measures differentiate Parkinson's disease (PD) patients from healthy controls (HC). *Front. Aging Neurosci.* **9**, 3 (2017).
- Siderowf, A. & Lang, A. E. Premotor Parkinson's disease: concepts and definitions. *Mov. Disord.* **27**, 608–616 (2012).
- Gaig, C. & Tolosa, E. When does Parkinson's disease begin? *Mov. Disord.* **24**, S656–S664 (2009).
- Liu, H., Wang, E. Q., Verhagen Metman, L. & Larson, C. R. Vocal responses to perturbations in voice auditory feedback in individuals with Parkinson's disease. *PLoS ONE* **7**, e33629 (2012).
- Chen, X. et al. Sensorimotor control of vocal pitch production in Parkinson's disease. *Brain Res.* **1527**, 99–107 (2013).
- Mollaei, F., Shiller, D. M. & Gracco, V. L. Sensorimotor adaptation of speech in Parkinson's disease. *Mov. Disord.* **28**, 1668–1674 (2013).
- Mollaei, F., Shiller, D. M., Baum, S. R. & Gracco, V. L. Sensorimotor control of vocal pitch and formant frequencies in Parkinson's disease. *Brain Res.* **1646**, 269–277 (2016).
- Sapir, S. Multiple factors are involved in the dysarthria associated with Parkinson's disease: a review with implications for clinical practice and research. *J. Speech Lang. Hear. Res.* **57**, 1330–1343 (2014).
- Huang, X. et al. The impact of Parkinson's disease on the cortical mechanisms that support auditory-motor integration for voice control. *Hum. Brain Mapp.* **37**, 4248–4261 (2016).
- Li, Y. et al. Neurobehavioral effects of LSVT® LOUD on auditory-vocal integration in Parkinson's disease: a preliminary study. *Front. Neurosci.* **15**, 624801 (2021).
- Dai, G. et al. Continuous theta burst stimulation over left supplementary motor area facilitates auditory-vocal integration in individuals with Parkinson's disease. *Front. Aging Neurosci.* **14**, 948696 (2022).

26. Shi, X., Wang, T., Wang, L., Liu, H. & Yan, N. Hybrid convolutional recurrent neural networks outperform CNN and RNN in task-state EEG detection for Parkinson's disease. In: *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. IEEE (2019).
27. van der Velden, B. H. M., Kuijff, H. J., Gilhuijs, K. G. A. & Viergever, M. A. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Med. Image Anal.* **79**, 102470 (2022).
28. Gunning, D. et al. XAI-Explainable artificial intelligence. *Sci. Robot* **4**, eaay7120 (2019).
29. Zhang, L., Wang, M., Liu, M. & Zhang, D. A survey on deep learning for neuroimaging-based brain disorder analysis. *Front. Neurosci.* **14**, 779 (2020).
30. Springenberg, J. T., Dosovitskiy, A., Brox, T. & Riedmiller, M. Striving for simplicity: The all convolutional net. *arXiv arXiv*, 1412.6806 (2014).
31. Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. *arXiv arXiv*, 1609.02907 (2016).
32. Wu, C., Wu, X.-J. & Kittler, J. Spatial residual layer and dense connection block enhanced spatial temporal graph convolutional network for skeleton-based action recognition. In: *Proceedings of the IEEE/CVF international conference on computer vision workshops* (2019).
33. Wu, B., Zhong, L., Li, H. & Ye, Y. Efficient complementary graph convolutional network without negative sampling for item recommendation. *Knowledge-Based Syst.* **256**, 109758 (2022).
34. Lawhern, V. J. et al. EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces. *J. Neural Eng.* **15**, 056013 (2018).
35. Yue, L. et al. Intention recognition from spatio-temporal representation of EEG signals. In: *Australasian Database Conference*. Springer (2021).
36. Vaswani, A. et al. Attention is all you need. *Adv. Neur. Inf. Process. Syst.* **30** (2017).
37. Yuvaraj, R., Acharya, U. R. & Hagiwara, Y. A novel Parkinson's Disease Diagnosis Index using higher-order spectra features in EEG signals. *Neural. Comput. Appl.* **30**, 1225–1235 (2018).
38. Shah, S. A. A., Zhang, L. & Bais, A. Dynamical system based compact deep hybrid network for classification of Parkinson disease related EEG signals. *Neural Netw.* **130**, 75–84 (2020).
39. Lee, S., Hussein, R., Ward, R., Jane Wang, Z. & McKeown, M. J. A convolutional-recurrent neural network approach to resting-state EEG classification in Parkinson's disease. *J. Neurosci. Methods* **361**, 109282 (2021).
40. Hassin-Baer, S. et al. Identification of an early-stage Parkinson's disease neuro-marker using event-related potentials, brain network analytics and machine-learning. *PLoS ONE* **17**, e0261947 (2022).
41. Guenther, F. H., Ghosh, S. S. & Tourville, J. A. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang.* **96**, 280–301 (2006).
42. Bohland, J. W., Bullock, D. & Guenther, F. H. Neural representations and mechanisms for the performance of simple speech sequences. *J. Cogn. Neurosci.* **22**, 1504–1529 (2010).
43. Chang, E. F., Niziolek, C. A., Knight, R. T., Nagarajan, S. S. & Houde, J. F. Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proc. Natl Acad. Sci. USA* **110**, 2653–2658 (2013).
44. Kort, N. S., Cuesta, P., Houde, J. F. & Nagarajan, S. S. Bihemispheric network dynamics coordinating vocal feedback control. *Hum. Brain Mapp.* **37**, 1474–1485 (2016).
45. Parkinson, A. L. et al. Understanding the neural mechanisms involved in sensory control of voice production. *NeuroImage* **61**, 314–322 (2012).
46. Behroozmand, R. et al. Sensory-motor networks involved in speech production and motor control: an fMRI study. *NeuroImage* **109**, 418–428 (2015).
47. Tait, L. et al. EEG microstate complexity for aiding early diagnosis of Alzheimer's disease. *Sci. Rep.* **10**, 17627 (2020).
48. Jouen, A. L., Lancheros, M. & Laganaro, M. Microstate ERP analyses to pinpoint the articulatory onset in speech production. *Brain Topogr.* **34**, 29–40 (2021).
49. Kindler, J., Hubl, D., Strik, W. K., Dierks, T. & Koenig, T. Resting-state EEG in schizophrenia: auditory verbal hallucinations are related to shortening of specific microstates. *Clin. Neurophysiol.* **122**, 1179–1182 (2011).
50. Costa, T. D. C. et al. Are the EEG microstates correlated with motor and non-motor parameters in patients with Parkinson's disease? *Neurophysiol. Clin.* **53**, 102839 (2023).
51. Guo, Z. et al. Top-down modulation of auditory-motor integration during speech production: the role of working memory. *J. Neurosci.* **37**, 10323–10333 (2017).
52. Liu, D. et al. Top-down inhibitory mechanisms underlying auditory-motor integration for voice control: evidence by TMS. *Cereb. Cortex* **30**, 4515–4527 (2020).
53. Li, T. et al. The left inferior frontal gyrus is causally linked to vocal feedback control: evidence from high-definition transcranial alternating current stimulation. *Cereb. Cortex* **33**, 5625–5635 (2023).
54. Brabenec, L. et al. Non-invasive brain stimulation for speech in Parkinson's disease: a randomized controlled trial. *Brain Stimul.* **14**, 571–578 (2021).
55. Brabenec, L. et al. Non-invasive stimulation of the auditory feedback area for improved articulation in Parkinson's disease. *Parkinsonism. Relat. Disord.* **61**, 187–192 (2019).
56. Wang, S., Zhang, P., Wang, H., Yu, H. & Zhang, F. Detecting shilling groups in online recommender systems based on graph convolutional network. *Inf. Process. Manag.* **59**, 103031 (2022).
57. Wang, X. et al. Traffic flow prediction via spatial temporal graph neural network. In: *Proceedings of the web conference 2020* (2020).
58. Parmiggiani, A., Liu, D., Psota, E., Fitzgerald, R. & Norton, T. Don't get lost in the crowd: graph convolutional network for online animal tracking in dense groups. *Comput. Electron. Agric.* **212**, 108038 (2023).
59. Zhang, X. et al. A survey on deep learning-based non-invasive brain signals: recent advances and new frontiers. *J. Neural Eng.* **18**, 031002 (2021).
60. Aliakbarhosseinabadi, S., Kamavuako, E. N., Jiang, N., Farina, D. & Mrachacz-Kersting, N. Online adaptive synchronous BCI system with attention variations. *Brain-Computer Interface Research: A State-of-the-Art Summary* **7**, 31–41 (2019).
61. Hughes, A. J., Daniel, S. E., Kilford, L. & Lees, A. J. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J. Neurol. Neurosurg. Psychiatr.* **55**, 181–184 (1992).
62. Chen, S. H., Liu, H., Xu, Y. & Larson, C. R. Voice F0 responses to pitch-shifted voice feedback during English speech. *J. Acoust. Soc. Am.* **121**, 1157–1163 (2007).
63. Lai, M., Demuru, M., Hillebrand, A. & Fraschini, M. A comparison between scalp- and source-reconstructed EEG networks. *Sci. Rep.* **8**, 12269 (2018).
64. Toth, B. et al. EEG network connectivity changes in mild cognitive impairment - Preliminary results. *Int. J. Psychophysiol.* **92**, 1–7 (2014).
65. Zhang, S., Tong, H. H., Tang, J., Xu, J. J. & Fan, W. Incomplete Network Alignment: Problem Definitions and Fast Solutions. *ACM Trans. Knowl. Discov. Data* **14**, 1–26 (2020).
66. Kuramochi, M. & Karypis, G. Finding frequent patterns in a large sparse graph. *Data. Min. Knowl. Disc.* **11**, 243–271 (2005).
67. Simonyan, K., Vedaldi, A. & Zisserman, A. Deep inside convolutional networks: Visualising image classification models and saliency maps. *arXiv arXiv*, 1312.6034 (2013).
68. Xia, F. et al. Graph learning: a survey. *IEEE Trans. Artif. Intell.* **2**, 109–127 (2021).
69. Liu, Y. et al. Graph self-supervised learning: a survey. *IEEE Trans. Knowl. Data Eng.* **35**, 5879–5900 (2022).
70. Koutra, D., Tong, H. & Lubensky, D. Big-align: Fast bipartite graph alignment. In: *2013 IEEE 13th international conference on data mining*. IEEE (2013).
71. Zhang, J. & Philip, S. Y. Multiple anonymized social networks alignment. In: *2015 IEEE International Conference on Data Mining*. IEEE (2015).
72. Singh, R., Xu, J. & Berger, B. Pairwise global alignment of protein interaction networks by matching neighborhood topology. In: *Annual international conference on research in computational molecular biology*. Springer (2007).
73. Liao, C. S., Lu, K., Baym, M., Singh, R. & Berger, B. IsoRankN: spectral methods for global alignment of multiple protein networks. *Bioinformatics* **25**, i253–i258 (2009).
74. Defferrard, M., Bresson, X. & Vandergheynst, P. Convolutional neural networks on graphs with fast localized spectral filtering. *Adv. Neural Inf. Process. Syst.* **29**, (2016).
75. Selvaraju, R. R. et al. Grad-cam: visual explanations from deep networks via gradient-based localization. In: *Proceedings of the IEEE international conference on computer vision*. IEEE (2017).
76. Michel, C. M. & Koenig, T. EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. *NeuroImage* **180**, 577–593 (2018).
77. Pascual-Marqui, R. D., Michel, C. M. & Lehmann, D. Segmentation of brain electrical activity into microstates: model estimation and validation. *IEEE Trans. Biomed. Eng.* **42**, 658–665 (1995).

ACKNOWLEDGEMENTS

This study was funded by grants from the National Key R&D Program of China (2018YFA0701405), National Natural Science Foundation of China (82172528, 81972147, 62271477, 82372556, 82302859, 82102648), Guangdong Basic and Applied Basic Research Foundation (2022A1515011203), Guangdong Province Science and Technology Planning Project (2017A050501014), Guangzhou Science and Technology Programme (201604020115), Shenzhen Science and Technology Program (KQTD20200820113106007), Shenzhen Science and Technology Program Grant Award (JCYJ20210324115810030, JCYJ20220818101217037, JCYJ20220818101411025), and The Science and Technology Planning Project of Guangdong Province (2023B1212060018).

AUTHOR CONTRIBUTIONS

X.C., N.Y. and H.L. conceived, designed, and coordinated the study. G.D., J.L., X.Z., X.H., Y.L., M.T. and X.C. recruited patients and completed data extraction. S.Z., G.D., J.L., X.Z. and X.H. contributed to performing the data analysis. S.Z., L.W., P.F., N.Y. and H.L. contributed to interpreting the results. S.Z., N.Y. and H.L. drafted and revised the manuscript. All authors approved the final version of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41746-023-00983-9>.

Correspondence and requests for materials should be addressed to Xi Chen, Nan Yan or Hanjun Liu.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024