

Supplementary Material

1. PARAMETER SETTINGS OF THE COMPETING MODELS

a) BrainGNN [1] BrainGNN is a cutting-edge graph neural network (GNN) framework that integrates innovative Ra-GConv layers and ROI-selection pooling (R-pool) layers, which effectively utilizes the topological and functional information extracted from fMRI data and facilitating the identification of predictive ROIs.

b) M-GAT-BC [2]: The M-GAT-BC framework is a neural network model that utilizes bilinear convolution (BC) for diagnosing SZ. It consists of two sections: the M-GAT section and the BC section. The M-GAT section is responsible for learning multiple node interaction features on each graph topology, while the BC section focuses on learning the spatial connectivity features of the brain network for predicting the disease.

c) SGP-SL [3], SGP-SL is an EEG-based MDD detection model named Self-attention Graph Pooling with Soft L label (SGP-SL). Specifically, they explore the local and global connections among EEG channels to construct an EEG-based graph in advance. By leveraging multiple self-attention graph pooling modules, the constructed graph is then gradually refined, followed by graph pooling, to aggregate information from less-important nodes to more-important ones. In this way, the feature representation with better discriminability can be learned from EEG signals. In addition, the soft label strategy is also adopted to build the loss function, aiming to further enhance the feature discriminability.

d) Cross-GNN[4] Cross-GNN is a Cross-modal Graph Neural Network that captures inter-modal dependencies through dynamic graph learning and mutual learning. Specifically, the inter-modal representations are attentively coupled into a compositional space for reasoning inter-modal dependencies. Additionally, they investigate mutual learning in explicit and implicit ways: (1) Cross-modal representations are obtained by cross-embedding explicitly based on the inter-modal correspondence matrix. (2) they propose a cross-modal distillation method to implicitly regularize latent representations with cross-modal semantic contexts. We carry out statistical analysis on the attentively learned correspondence matrices to evaluate inter-modal relationships for associating disease biomarkers. Our extensive experiments on three datasets demonstrate the superiority of our proposed method for disease diagnosis with promising prediction performance and multi-modal connectome biomarker location.

e) TAN [5]: TAN introduces a triple network to extract discriminative information from the high-order representation feature space obtained from multi-modal data. It incorporates self-attention to dynamically estimate the importance of brain regions and utilizes the cross-attention mechanism to extract complementary information from fMRI and DTI.

f) RH-BrainFS [6]: In this study, aiming to address the challenge of regional heterogeneity in multimodal brain networks, a novel Regional Heterogeneous Multimodal Brain Networks Fusion Strategy is proposed. This strategy incorporates a brain subgraph networks module to extract regional characteristics from brain networks and employs a novel transformer-based fusion bottleneck module to mitigate regional heterogeneity between structural connectivity (SC) and functional connectivity (FC).

g) MCRLN[4] In this study, a multi-connectivity representation learning framework is developed to integrate topological representations from structural connectivity, functional connectivity, and dynamic functional connectivities for the automatic diagnosis of Major Depressive Disorder (MDD). Specifically, structural, static functional, and dynamic functional graphs are first computed from diffusion magnetic resonance imaging (dMRI) and resting-state functional magnetic resonance imaging (rsfMRI). Then, a novel Multi-Connectivity Representation Learning Network (MCRLN) is proposed to integrate these multiple graphs using structural-functional fusion and static-dynamic fusion modules. The Structural-Functional Fusion (SFF) module is innovatively designed to decouple graph convolution, enabling the capture of both modality-specific and modality-shared features for precise brain region representation. To further integrate static and dynamic functional graphs, a novel Static-Dynamic Fusion (SDF) module is developed to transfer important connections from static graphs to dynamic graphs through attention mechanisms. Finally, the proposed approach's performance is thoroughly evaluated on large clinical datasets, demonstrating its effectiveness in classifying MDD patients.

Table S1. Parameter settings of competing models.

Method	Hyper-parameter	Range
BrainGNN	conv layers	2
	pooling layers	2
	learning rate	0.001
	Others	Default
M-GAT-BC	output feature dimension of node	$h = N/2 = 60$.
	Learning rate (LR)	0.001
	numbers of channels	$Q = 32$ $C = 64$
	Optimizer	Adam ($\beta_1 = 0.9, \beta_2 = 0.999, \epsilon = 1e-8$)
SGP-SL	λ	0.1
	γ	0.01
	ϵ	0.1
	Others	Default
Cross-GNN	layers of bilateral graph convolution	grid search from 1 to 4.
	multi-layer perception	3-layer
	learning rate	$3e-4$
	weight decay	$5e-5$
TAN	Others	Default
	Optimizer	Adam
	LR	$5e-4$
	α ,	0.8
	β	0.5
	p	0.5
RH-BrainFS	other	Default
	step-size	50
	gamma	0.8
	learning rate	$5e-4$
	dropout rate	0.3
	k-hop in the subgraph sampling	1
	the number of bottlenecks N_b	3
	the number of attention heads in the Transformer	3
	total number of network layers	2
MCRLN	Optimizer	Adam
	LR	$5e-4$
	weight decay	$5e-5$
	tradeoff parameter	$\{10^{-5}, 10^{-4}, \dots, 10^5\}$
MPP-GCN	Others	Default
	Optimizer	Adam ($\beta_1 = 0.9$ $\beta_2 = 0.999, \epsilon = 1e-8$)
	LR	Log-uniform (min = $1e-5$, max = $1e-2$)
	Input/output dimensions	Discrete-uniform (min = 5, max = 16, step = 1)
	Others	Default

h) MMP-GCN [5]:MMP-GCN introduces a brain connectivity network that combines resting-state functional magnetic resonance imaging (fMRI) and diffusion tensor imaging (DTI) to diagnose Alzheimer’s Disease. It proposes three mechanisms within the graph convolutional network (GCN) to enhance performance. The GCN parameters for all strategies in this paper are fixed and chosen based on previous work [7].

2. PARAMETER SETTINGS OF THE PROPOSED MODEL

Below is an English summary of the key model hyperparameters and training settings used in FE-NET, organized into three brief paragraphs, followed by a consolidated table.

The GAN-based fMRI-EEG hypergraph generation (FEH) module uses an Interactive Hyper-edge Neurons (IHEN) generator composed of $L = 5$ stacked IHEN layers. Within each IHEN block, node and hyperedge feature transforms are controlled by a balance hyperparameter $\eta = 0.6$, which weights self-updates versus cross-updates between nodes and hyperedges.

In the dynamic fMRI-EEG hypergraph embedding (FED) module, we treat the diffusion step as a random-walk teleportation process with two teleportation probabilities: $\alpha_v = 0.05$ for vertex retention and $\alpha_e = 0.90$ for hyperedge retention. The control step is implemented via a dual-stream control function (DSCF) combining a fully-connected layer, an LSTM, and an attention block; no additional knobs beyond α_v and α_e are tuned here.

All components are learned end-to-end using 5-fold cross-validation. We optimize with Adam at an initial learning rate of 1×10^{-2} , weight decay of 5×10^{-4} , and a dropout rate of 0.20 applied to all hidden layers. GAN-style adversarial training alternates between the generator (FEH) and a standard MLP discriminator.

Table S2. Summary of Hyperparameters and Settings of proposed model

Module/Stage	Parameter	Value
FEH generator depth	Number of IHEN layers	5
FEH cross-update weight	η	0.60
FED diffusion teleportation	α_v (vertex)	0.05
FED diffusion teleportation	α_e (hyperedge)	0.90
Optimizer	Learning rate	1×10^{-2}
Regularization	Weight decay	5×10^{-4}
Regularization	Dropout rate	0.20
Validation scheme	Cross-validation	5-fold
Discriminator architecture	Model	3-layer MLP

3. INPUT AND OUTPUT PARAMETERS FOR DUAL-STREAM CONTROL FUNCTION (DSCF) LAYER

Table S3. Configurations of F, L and A

Component	Layer	Operation	Output
F	Fully Connected	Softmax	2
L	Input (Sequence)		50×2048
	B-LSTM	Cell number is 128	50×256
	B-LSTM	Cell number is 64	50×128
A	Input	Concatenate	256×128
	Attention	Self-Attention	128

4. DATA ALIGNMENT AND NODE CORRESPONDENCE

1. Node Alignment Strategy

We confirm explicit 1:1 anatomical correspondence** between EEG and fMRI nodes via standardized neuroimaging pipelines: fMRI parcellation: All subjects use the AAL-90 atlas (90 cortical/subcortical ROIs). EEG processing: Full 62-electrode data undergo source localization (sLORETA) to cortical voxels. Voxel-level signals are spatially aggregated to the same 90 AAL ROIs as fMRI.

2. EEG Source Localization Details The pipeline is:

Forward modeling: Boundary element method (BEM) with standard head model. Inverse solution: sLORETA for source reconstruction. ROI aggregation: Voxel-level source signals averaged within each AAL-90 region. EC computation: Partial directed coherence (PDC) between ROI-level time series. EEG signals were source-localized to AAL-90 ROIs via sLORETA. ROI-level bandpower features and PDC-based EC matrices were then computed.

5. SUPPLEMENTARY EXPERIMENT

Table S4. Comparing FE-NET's performance with traditional methods

TAP-WST	Method	Acc	Pre	Spe	TMT-LPS	Acc	Pre	Spe
SA-TASK	FE-NET	85.29±2.57	83.46±3.19	80.83±4.05	CF-TASK	79.18±3.47	75.29±4.83	74.38±5.62
	FE-NET(RNN)	76.45±3.82	75.63±2.47	71.91±2.73		71.57±3.64	67.27±4.18	65.84±3.59
	FE-NET(LSTM)	77.35±2.67	78.17±5.83	75.82±6.45		68.46±8.53	65.38±5.29	66.49±5.84
	FE-NET(GRU)	79.90±2.90	76.45±4.00	73.80±5.20		73.30±3.50	70.10±4.50	68.60±6.10
VI-TASK	FE-NET	82.19±2.84	80.47±3.85	78.53±4.73	LI-TASK	77.38±3.47	73.29±4.85	72.84±5.47
	FE-NET(RNN)	76.29±3.85	75.38±2.84	72.62±2.47		72.47±3.29	68.53±4.18	65.29±3.85
	FE-NET(LSTM)	78.84±2.73	78.62±5.29	74.47±6.18		69.53±8.47	66.38±5.84	66.29±5.47
	FE-NET(GRU)	79.60±2.80	77.20±4.10	73.50±5.30		70.90±4.00	69.50±4.90	67.30±6.00

We have attempted to use traditional methods to model the process of transitioning from $x(t_k)$ to $x(t_{k+1})$. As shown in Table S4, the performance of FE-NET (RNN, LSTM, GRU) is not as effective as FE-NET.

Generalization capability evaluation: To further evaluate our generalization capability, we conducted additional benchmarking on the synchronized CN-EPFL datasets. In Figure S1(upper), all methods were trained on LEMON and tested on CN-EPFL, while in Figure S1(lower), all models were both trained and tested on CN-EPFL. The results demonstrate that our model still exhibits significant advantages, whether generalizing to task-related datasets or directly utilizing synchronized datasets, further validating its superiority.

As shown in Figure S2 (top), the ablation study of OFEI and IHEN in FEH is conducted by evaluating the average accuracy in four downstream tasks. The variants of FED that exclusively employ OFEI or IHEN are referred to as "FEH(OFEI)" and "FEH(IHEN)", respectively. The performance of FEH outperforms "FEH(OFEI)" and "FEH(IHEN)", demonstrating its effectiveness. Additionally, as shown in Figure S2 (bottom), the efficacy of the control and diffusion processes within the FED is validated. The variants of FED that employ exclusively control processes or

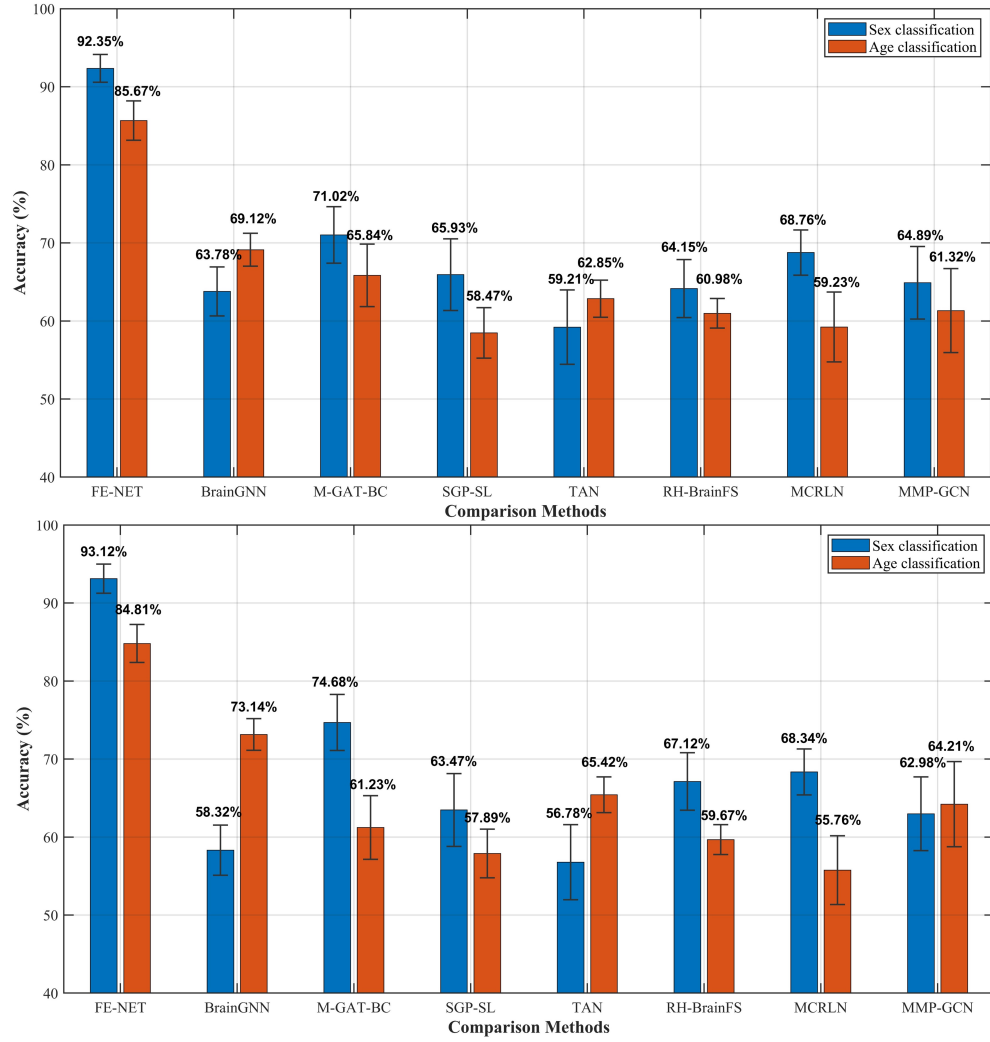


Fig. S1. Benchmarking on the synchronized CN-EPFL datasets.

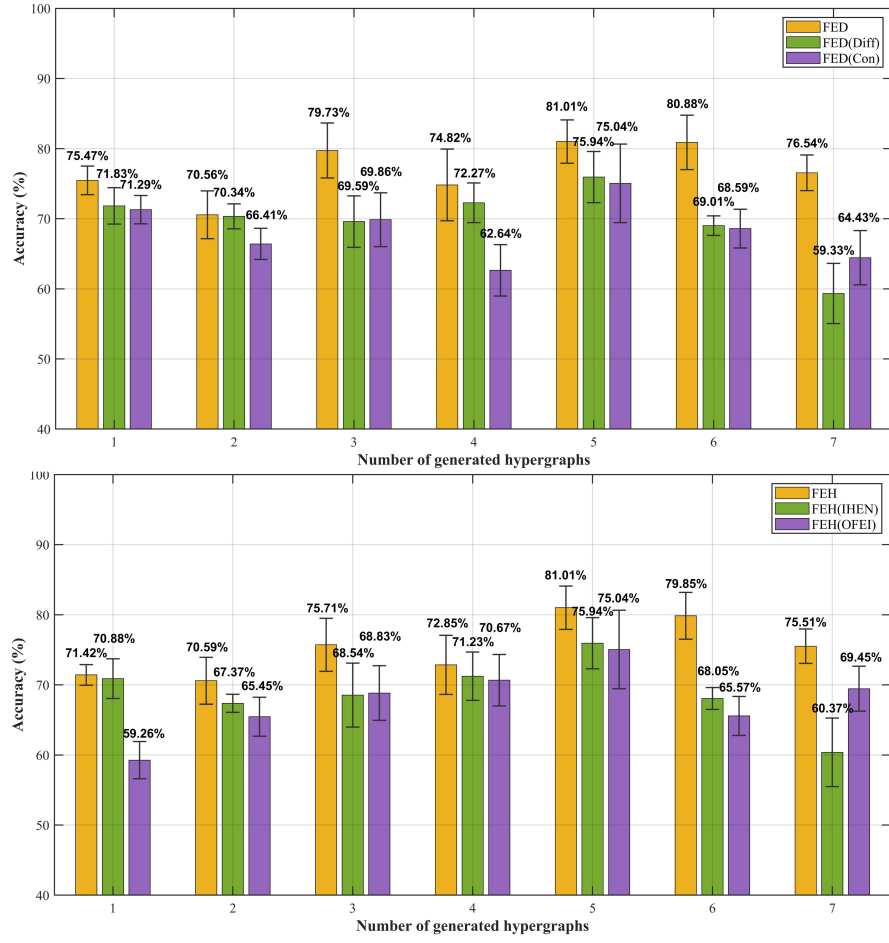


Fig. S2. FED, and FEH modules analysis.

diffusion processes are denoted as "FED(Con)" and "FED(Diff)" respectively. A comparative analysis of the two processes reveals that the inclusion of only the control process or diffusion process, which relies solely on initial vertex features and neglects the hypergraph structure, results in inferior outcomes. A comparative analysis of the two processes reveals that the inclusion of only the control process or diffusion process, which relies solely on initial vertex features and neglects the hypergraph structure, results in inferior outcomes.

6. TIME COMPLEXITY ANALYSIS

FEH Module (GAN-based Hypergraph Generation)

- **OFEI Algorithm:** $O(n^2m)$, where n = number of ROIs (typically 90–120), m = hyperedges per subject.
- **IHEN Layers:** $O(L(N_v d^2 + N_e k^2))$, where L = 5 layers, N_v = vertices, N_e = hyperedges, d = feature dimension (256), k = hyperedge cardinality.

FED Module (Neural ODE Embedding)

- **Control Step:** $O(T(N_v + N_e)d^2)$, where $T \approx 50$ integration steps. The dual-stream LSTM with attention dominates.
- **Diffusion Step:** $O(N_v^2 + N_e^2)$ per step due to hypergraph Laplacian operations. Sparse matrix optimizations reduce practical cost.

GPU Memory Usage

Table S5. Approximate VRAM Consumption on NVIDIA A6000

Component	VRAM Consumption*	Notes
FEH Generator (IHEN \times 5)	4–8 GB	Batch size 8–16
GAN Discriminator	6–8 GB	MLP with 3×512 layers
FED ODE Solver	7–9 GB	Adjoint method + intermediates
Hypergraph Data (N=100)	3–5 GB	Storing $M = 1$ K hyperedges

*Measured during peak training/inference.

FE-NET Runtime on LEMON Using 1 NVIDIA A6000

Table S6. FE-NET Training Performance

Metric	Value
Batch Size	16 subjects
Epoch Time	18.1 minutes
Total Training (200 epochs)	36.1 hours
Throughput	6.3 samples/minute

Training Performance

Inference Performance At the same time, we provided a comparison of computational resources with other SOTA methods in the Table S8. Although FE-NET requires more computational resources than existing SOTA methods, The downstream tasks of fMRI and EEG imaging are not test-intensive, the slightly longer inference time after model training is completed will not interfere with the actual application of the model.

In fact, when the batch size is adjusted to 8, our model can run perfectly on a single 2080ti. Existing larger fMRI datasets are typically at the thousand-level scale, and even if thousand-level fMRI-EEG datasets emerge in the future, four 2080ti GPUs can handle them perfectly. Considering

Table S7. FE-NET Inference Performance

Metric	Value
Single-subject Inference	3.8s \pm 0.12s
Batch Inference (16 subjects)	30.8s \pm 1.92s
VRAM Usage (Inference)	22.4 GB

the cost of acquiring fMRI and EEG data, four NVIDIA 2080ti GPUs is a small price to pay. Given the considerable potential for future improvements in GPU performance, our model remains well-suited for large-scale datasets.

Table S8. Computational Resource Comparison using NVIDIA A6000

Method	Batch Size	VRAM Usage (GB)	Batch Inference Time (s)	Training Time (h)	Single-subject Inference Time (s)	Peak GPU Utilization (%)
GCN [8]	16	4.3 \pm 0.2	3.36 \pm 0.32	2.8	0.21 \pm 0.02	10.1 \pm 2.1
GAT [9]	16	5.5 \pm 0.3	5.44 \pm 0.48	3.5	0.34 \pm 0.03	16.3 \pm 1.8
GIN [10]	16	6.1 \pm 0.3	4.96 \pm 0.48	4.2	0.31 \pm 0.03	18.5 \pm 3.2
BrainGNN [1]	16	10.7 \pm 0.5	22.72 \pm 1.60	6.8	1.42 \pm 0.10	50.2 \pm 2.5
M-GAT-BC [2]	16	15.2 \pm 0.6	19.36 \pm 1.60	8.5	1.21 \pm 0.10	53.7 \pm 2.0
Cross-GNN [4]	16	18.3 \pm 0.7	51.68 \pm 3.20	9.8	3.23 \pm 0.20	66.4 \pm 1.7
MMP-GCN [11]	16	20.1 \pm 0.8	44.96 \pm 2.88	10.5	2.81 \pm 0.18	59.3 \pm 1.3
FE-NET	16	22.4 \pm 0.9	60.80 \pm 1.92	36.1	3.8 \pm 0.12	72.4 \pm 1.1

7. DETAILS ON THE DATASET AND DATA PREPROCESSING.

A. CN-EPFL dataset

Twenty healthy, right-handed volunteers (mean age $a \approx b$ 24.6 years) with normal or corrected vision performed two conditions of a speeded numerosity task while EEG and fMRI were recorded. In the active (“committed”) condition, participants decided which of two 60 ms dot arrays (total of 100 dots) contained more dots, responded within 500 ms, and then rated their confidence on a continuous scale from “sure error” (0) to “sure correct” (1). In the observation condition, they watched a computer-controlled hand make the same choices (replays of active trials, shuffled) and then rated confidence in the observed decision. Each condition comprised 144 trials per subject, with difficulty adaptively titrated to yield 71 % accuracy.

Behavioral preprocessing excluded trials with missing first-order (2.0 %) or confidence responses (2.9 %), RTs faster than 200 ms or slower than 500 ms (13.1 %), and any observation trials with unintended button presses (12.6 %). After exclusions, each subject contributed about 119 ± 5 trials per condition for analysis of choice accuracy, response times, and confidence ratings.

EEG data were acquired from 63 channels at 5 kHz simultaneous with fMRI. Gradient and cardioballistic artifacts were removed via sliding-window template subtraction and heartbeat-triggered ICA. The continuous signal was band-pass filtered (1–10 Hz), epoched from –200 to +600 ms around the first-order response (button press or hand onset), rereferenced to the common average, and subjected to ICA and dipolar source fitting. Nonbrain or artifactual components were discarded, yielding on average eight brain-related components per subject for single-trial analyses.

fMRI data (3 T Siemens Prisma, multiband EPI, TR = 1,280 ms, TE = 31 ms, 2 mm isotropic voxels) were preprocessed in SPM12. Functional volumes were realigned, resliced, and normalized to MNI space via DARTEL; then spatially smoothed with a 5 mm FWHM Gaussian kernel. First-level models included regressors for each condition and parametric modulators (confidence, response time, stimulus evidence), along with motion parameters and excluded-trial regressors; group-level contrasts were FWE-corrected at $P < 0.05$ (cluster level, voxel threshold $P < 0.001$).

B. LEMON dataset

This dataset offers a large, openly shared collection of simultaneous electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) recordings gathered from 22 healthy adults (ages 23–51) at the Nathan S. Kline Institute. Participants viewed a range of visual and naturalistic stimuli—flickering checkerboards, rest with eyes open, the computer-generated animation *Inscapes*, and several short “real-world” movies—both outside the scanner and during simultaneous EEG–fMRI acquisition. Alongside EEG and BOLD signals, the study also captured eye-tracking, electrocardiography (ECG), respiration, and behavioral/cognitive measures, providing a rich multimodal resource for exploring the spatiotemporal dynamics of human brain activity.

Each subject attended two sessions, each including an “Outside” EEG-only checkerboard run, a “Scanner OFF” EEG run inside the MR bore, and a “Scanner ON” run combining EEG and fMRI. Naturalistic viewing was split across multiple short movie clips (e.g. *The Present*, *Despicable Me* in English and Hungarian, and three monkey-video segments) interleaved with block-design tasks. All data are organized according to the Brain Imaging Data Structure (BIDS), with raw files (EEG .set/.vhdr/.vmrk, fMRI NIfTI, eye-tracking TSV) accompanied by metadata sidecars and behavioral logs. Preprocessed derivatives are available in parallel folder hierarchies.

EEG preprocessing was carried out in EEGLAB, using a graduated pipeline tuned to acquisition setting. “Outside” and “Scanner OFF” data underwent bandpass filtering (0.5–70 Hz), average referencing (excluding ECG/EOG channels), and, where applicable, pulse-artifact (ballistocardiogram) removal via template subtraction. “Scanner ON” data then received gradient-artifact correction with the FMRIB plugin’s optimal basis set subtraction, followed by PCA-based residual cleanup, pulse-artifact removal, re-filtering, and re-referencing. Automated quality checks flagged “good” channels and trials, and independent-component analysis (ICA) identified brain-source components.

fMRI preprocessing used the Connectome Computation System framework: structural T1w images were skull-stripped (BET+FreeSurfer), segmented, and normalized to MNI space (FLIRT, MCFLIRT). Functional runs discarded the first five volumes, then applied despiking, slice-time and motion correction, brain masking, boundary-based registration, and nuisance regression (Friston-24 motion parameters, CSF/WM signals, optional global-signal regression). Users can choose temporal filtering (0.01–0.1 Hz) and spatial smoothing (6 mm FWHM). Finally, time series were extracted from 400 cortical parcels (Schaefer atlas) for downstream analyses.

Table S9. Abbreviations of different ROIs and brain regions

AAL Regions	Abbreviation	AAL Regions	Abbreviation
Precentral gyrus	PreCG	Lingual gyrus	LING
Superior frontal gyrus, dorsolateral	SFGdor	Superior occipital gyrus	SOG
Superior frontal gyrus, orbital part	ORBsup	Middle occipital gyrus	MOG
Middle frontal gyrus	MFG	Inferior occipital gyrus	IOG
Middle frontal gyrus, orbital part	ORBmid	Fusiform gyrus	FFG
Inferior frontal gyrus, opercular part	IFGoperc	Postcentral gyrus	PoCG
Inferior frontal gyrus, triangular part	IFGoperc	Superior parietal gyrus	SPG
Inferior frontal gyrus, orbital part	ORBinf	Inferior parietal, but supramarginal and angular gyri	IPL
Rolandic operculum	ROL	Supramarginal gyrus	SMG
Supplementary motor area	SMA	Angular gyrus	ANG
Olfactory cortex	OLF	Precuneus	PCUN
Superior frontal gyrus, medial	SFGmed	Paracentral lobule	PCL
Superior frontal gyrus, medial orbital	ORBsupmed	Caudate nucleus	CAU
Gyrus rectus	REC	Lenticular nucleus, putamen	PUT
Insula	INS	Lenticular nucleus, pallidum	PAL
Anterior cingulate and paracingulate gyri	ACG	Thalamus	THA
Median cingulate and paracingulate gyri	DCG	Heschl gyrus	HES
Posterior cingulate gyrus	PCG	Superior temporal gyrus	STG
Hippocampus	HIP	Temporal pole: superior temporal gyrus	TPOsup
Parahippocampal gyrus	PHG	Middle temporal gyrus	MTG
Amygdala	AMYG	Temporal pole: middle temporal gyrus	TPOmid
Calcarine fissure and surrounding cortex	CAL	Inferior temporal gyrus	ITG
Brain Regions	Abbreviation	Brain Regions	Abbreviation
default mode network	DMN	frontoparietal network	FPN
limbic network	LN	ventral attention network	VAN
sensorimotor network	SMN	dorsal attention network	DAN
visual network	VN	subcortical system	SUB

REFERENCES

1. X. Li, Y. Zhou, N. Dvornek, M. Zhang, S. Gao, J. Zhuang, D. Scheinost, L. H. Staib, P. Ventola, and J. S. Duncan, "Braingnn: Interpretable brain graph neural network for fmri analysis," *Med. Image Analysis* **74**, 102233 (2021).
2. R. Yu, C. Pan, X. Fei, M. Chen, and D. Shen, "Multi-graph attention networks with bilinear convolution for diagnosis of schizophrenia," *IEEE J. Biomed. Heal. Informatics* **27**, 1443–1454 (2023).
3. T. Chen, Y. Guo, S. Hao, and R. Hong, "Exploring self-attention graph pooling with eeg-based topological structure and soft label for depression detection," *IEEE transactions on affective computing* **13**, 2106–2118 (2022).
4. Y. Yang, C. Ye, X. Guo, T. Wu, Y. Xiang, and T. Ma, "Mapping multi-modal brain connectome for brain disorder diagnosis via cross-modal mutual learning," *IEEE Transactions on Med. Imaging* (2023).
5. Q. Zhu, H. Wang, B. Xu, Z. Zhang, W. Shao, and D. Zhang, "Multimodal triplet attention network for brain disease diagnosis," *IEEE Transactions on Med. Imaging* **41**, 3884–3894 (2022).
6. H. Ye, Y. Zheng, Y. Li, K. Zhang, Y. Kong, and Y. Yuan, "Rh-brainfs: Regional heterogeneous multimodal brain networks fusion strategy," *Adv. Neural Inf. Process. Syst.* **36** (2024).
7. S. Parisot, S. I. Ktena, E. Ferrante, M. Lee, R. Guerrero, B. Glocker, and D. Rueckert, "Disease prediction using graph convolutional networks: application to autism spectrum disorder and alzheimer's disease," *Med. image analysis* **48**, 117–130 (2018).
8. T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907* (2016).
9. P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," *arXiv preprint arXiv:1710.10903* (2017).
10. K. Xu, W. Hu, J. Leskovec, and S. Jegelka, "How powerful are graph neural networks?" *arXiv preprint arXiv:1810.00826* (2018).
11. X. Song, F. Zhou, A. F. Frangi, J. Cao, X. Xiao, Y. Lei, T. Wang, and B. Lei, "Multicenter and multichannel pooling gcn for early ad diagnosis based on dual-modality fused brain network," *IEEE Transactions on Med. Imaging* **42**, 354–367 (2023).