UniFCHarm: Unified Prompt-Guided Harmonization of fMRI Functional Connectivity Across Scanners and Atlases

Xin Lin[1†], Yuxiao Liu[1†], Yulin Wang[1], Lin Teng[1], Long Yang[1], Shilun Zhao[1], Kai Zhang[1], Zhiming Cui[1], and Dinggang Shen[1,2,3(✉)]

[1] School of Biomedical Engineering & State Key Laboratory of Advanced Medical Materials and Devices, ShanghaiTech University, Shanghai 201210, China
[2] Shanghai United Imaging Intelligence Co., Ltd., Shanghai 200230, China
[3] Shanghai Clinical Research and Trial Center, Shanghai 201210, China
Dinggang.Shen@gmail.com

Abstract. Multi-site functional MRI (fMRI) is essential for learning generalizable functional connectivity (FC) biomarkers, yet scanner- and protocol-specific variability shifts FC distributions and degrades cross-site generalization, necessitating effective harmonization. Existing methods often rely on simple many-to-one global mappings toward a single pre-defined target site and are also difficult to transfer across atlases. To address these issues, we propose **UniFCHarm**, a unified prompt-guided framework for multi-scanner and multi-atlas FC harmonization. UniFCHarm first pretrains a GraphCLIP using 44,493 acquisition- and atlas-aware FC-prompt pairs to align them into a shared latent space, enabling harmonization toward a user-specified target condition instead of being constrained to a single reference site or atlas. It then performs coarse-to-fine harmonization by first correcting subnetwork-to-subnetwork connectivity and subsequently refining ROI-to-ROI connections, thereby mitigating over-correction. Evaluations on traveling-subject benchmark across multiple atlases demonstrate that UniFCHarm achieves the highest ground-truth agreement with MAE 0.033 ± 0.001 and reliably matches the target acquisition setting with accuracy $91.25 \pm 6.18\%$, outperforming the evaluated representative baselines. Our code is available at `https://github.com/yizhixiaolinya/Code-MICCAI2026_fMRI-Harm.git`.

Keywords: fMRI harmonization, functional connectivity, prompt conditioning, multi-scanner neuroimaging, generative modeling.

1 Introduction

Multi-site functional MRI (fMRI) studies are essential for learning generalizable functional-connectivity (FC) biomarkers for disease diagnosis and subtyping [12]. However, FC data acquired across sites exhibit heterogeneity driven by acquisition setting (including

[†] Equal contribution.

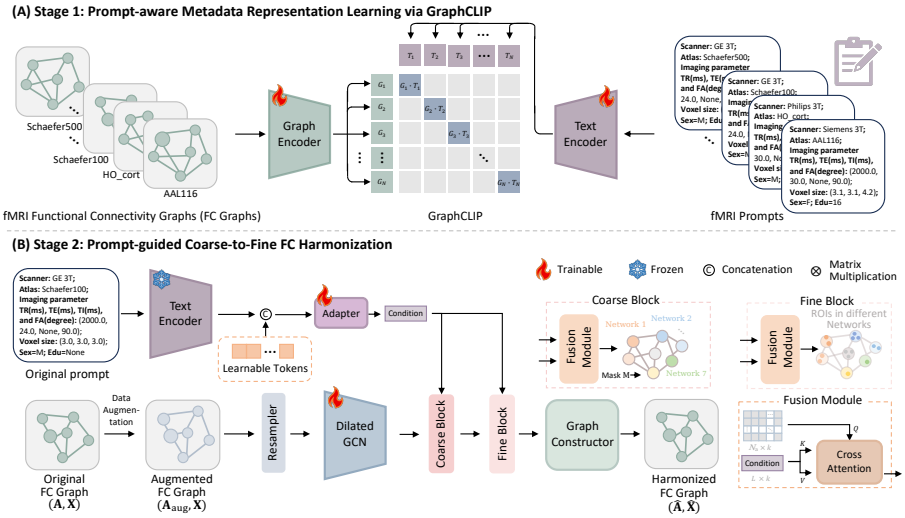

Fig. 1. Overview of the proposed UniFCHarm. (A) GraphCLIP pre-training for aligning multi-scanner FC with structured acquisition- and atlas-aware prompts. (B) Prompt-guided harmonization for translating an input FC toward a desired prompt-specified acquisition condition, enabling multi-scanner and multi-atlas harmonization.

scanner and acquisition parameter) differences, which affects FC distributions and degrades downstream analysis [26]. FC harmonization is therefore essential to mitigate such heterogeneity, while preserving subject-specific variability associated with subject characteristics and disease status [23].

To achieve FC harmonization, previous methods primarily employ linear parametric models (*e.g.*, ComBat [10]) that regress out scanner bias by adjusting FC means and covariances. While computationally efficient, these linear approaches are limited in modeling the complex relationship between high-dimensional FC data and acquisition-induced heterogeneity. To better capture non-linear effects, deep learning harmonization frameworks have emerged [20, 22]. Despite the effectiveness of these methods, they often remain constrained by a rigid many-to-one paradigm that maps all FC data toward a single, pre-defined target site [20], lacking the flexibility to handle diverse acquisition settings in realistic multi-site scenarios [22]. Moreover, most existing frameworks are atlas-dependent and cannot generalize across different brain atlases [26]. Beyond these flexibility constraints, many deep learning-based methods overlook the multi-scale nature of acquisition-induced heterogeneity [20, 8]. In practice, site effects can shift connectivity patterns between subnetworks while also introducing finer biases at the level of individual ROI-to-ROI connections. When both scales are handled by a single-step, non-selective global correction, the method lacks coarse-to-fine constraints to separate systematic site noise from subject-specific biological signals, increasing the risk of over-correction.

To address these limitations, we propose UniFCHarm, a unified prompt-guided framework for fMRI FC harmonization across acquisition settings and atlases. UniFCHarm

pretrains GraphCLIP to align structured prompts that encode acquisition settings and atlas specifications with FC graphs in a shared latent space, yielding prompt embeddings that condition harmonization toward a desired target setting. By conditioning on these embeddings rather than a fixed reference site, UniFCHarm performs flexible, prompt-directed harmonization that generalizes across arbitrary scanner-protocol-atlas combinations. To handle the multi-scale nature of site effects, UniFCHarm further adopts a hierarchical coarse-to-fine strategy by first harmonizing subnetwork-level connectivity and then refining ROI-level connections to correct residual biases, thereby constraining fine-grained harmonization within a coarse structural prior to reduce the risk of over-correction.

Our contributions are threefold. **First**, we introduce UniFCHarm, a prompt-guided harmonization framework that enables prompt-conditioned translation without mapping all data to a fixed reference site and with improved compatibility across atlases. **Second**, we propose a hierarchical coarse-to-fine harmonization strategy that first harmonizes subnetwork-to-subnetwork connectivity and then refines ROI-to-ROI connections to better handle multi-scale acquisition-induced heterogeneity while reducing the risk of over-correction. **Third**, extensive experiments on traveling-subject benchmark show that our method consistently outperforms representative baselines, yielding lower error to the ground truth with MAE 0.033 ± 0.001 and higher target-setting compliance with accuracy $91.25 \pm 6.18\%$.

2 Method

As depicted in Fig. 1, UniFCHarm is a two-stage framework that hierarchically harmonizes FCs using acquisition- and atlas-aware prompts. In Stage 1 (Fig. 1(A)), we pretrain GraphCLIP to align prompts with FC graphs in a shared latent space. In Stage 2 (Fig. 1(B)), the Prompt-guided Graph Harmonizer (PGH) conditions on these embeddings to perform hierarchical FC harmonization toward a user-specified target acquisition setting.

2.1 GraphCLIP Pre-training for FC-Prompt Alignment

We pretrain GraphCLIP with a contrastive objective. Specifically, we use a text encoder to embed the prompt and a graph encoder to extract FC representations, for aligning the two modalities in a shared latent space. The text encoder E_t is an eight-head Transformer that maps each structured prompt to a prompt embedding, and the graph encoder E_g is a three-layer GCN that maps each FC graph to a graph embedding.

Construction of FC-specific Text Prompt. FC distributions are influenced by the acquisition setting, and the atlas choice further shifts them by changing parcellation granularity and ROI definitions. To explicitly guide harmonization, we encode the acquisition setting and atlas specification into a structured prompt with a fixed template, as illustrated in Fig. 1(A) (right). The prompt includes scanner identity, TR, TE, TI, flip angle, voxel size, and FC atlas, allowing the target condition to be specified at inference.

Contrastive Learning-based Pre-training. To align FC graphs with structured prompts, we optimize E_g and E_t with a symmetric InfoNCE loss. For a mini-batch of B paired FC

graphs and prompts, we ℓ_2-normalize the embeddings to obtain $\mathbf{g}_i$ and $\mathbf{t}_i$. The objective maximizes the similarity of matched graph–prompt pairs and contrasts them against mismatched pairs:

$$\mathcal{L}_{\text{clip}} = -\frac{1}{2B}\sum_{i=1}^{B}\left[\log\frac{\exp(\text{sim}(\mathbf{g}_i,\mathbf{t}_i)/\tau_{\text{clip}})}{\sum_{j=1}^{B}\exp(\text{sim}(\mathbf{g}_i,\mathbf{t}_j)/\tau_{\text{clip}})} + \log\frac{\exp(\text{sim}(\mathbf{t}_i,\mathbf{g}_i)/\tau_{\text{clip}})}{\sum_{j=1}^{B}\exp(\text{sim}(\mathbf{t}_i,\mathbf{g}_j)/\tau_{\text{clip}})}\right], \tag{1}$$

where $\text{sim}(\cdot,\cdot)$ denotes cosine similarity and τ_{clip} is the InfoNCE temperature. After pre-training, the text encoder is frozen and used to produce target-prompt embeddings for Stage 2.

2.2 Prompt-guided Coarse-to-Fine FC Harmonization

Based on prompt embeddings in Stage 1, we design a PGH to perform hierarchical coarse-to-fine FC harmonization while mitigating over-correction. In Stage 2, we further train it in a self-supervised manner due to the lack of traveling subjects across different sites. We augment the original FC to obtain $\mathbf{A}_{\text{aug}}$ and train the PGH to recover the original FC conditioned on its prompt embedding extracted from the frozen text encoder in Stage 1. The original prompt is regarded as the target prompt during training. At inference, the prompt is a user-specified acquisition condition for target harmonization.

Data augmentation. We construct an acquisition-style augmentation $f_{\text{aug}}(\cdot)$ by matching global FC statistics to a reference sample. For each input FC $\mathbf{A}$, we randomly sample a reference FC $\mathbf{A}_{\text{ref}}$ acquired under a different acquisition setting. We then align the global edge-weight statistics of $\mathbf{A}$ to those of $\mathbf{A}_{\text{ref}}$ by matching their mean and standard deviation. We further introduce a scaling factor γ combined with tanh to induce non-linear shifts. This yields an augmented FC $\mathbf{A}_{\text{aug}} = f_{\text{aug}}(\mathbf{A};\mathbf{A}_{\text{ref}})$ that mimics an acquisition-induced domain shift. The model is then trained to remove this synthetic shift and recover the target condition specified by the prompt.

Resampler and dilated-GCN. To support multi-atlas harmonization within a single framework, and to resolve the dimensionality mismatch between atlas-dependent ROI features and the fixed parameterization of the graph encoder, we adopt a Flamingo-style Resampler [1]. For atlas a, we denote the number of ROIs by N_a, the input feature dimension by d, and the latent dimension by k. The Resampler applies cross-attention over the initial ROI features $\mathbf{X} \in \mathbb{R}^{N_a\times d}$ and outputs fixed-width ROI representations $\mathbf{Z} \in \mathbb{R}^{N_a\times k}$, yielding a consistent feature dimension across atlases. We then feed $\mathbf{Z}$ into a multi-scale dilated GCN with the augmented FC graph $\mathbf{A}_{\text{aug}}$ as the weighted adjacency:

$$\mathbf{H} = \text{GCN}(\mathbf{A}_{\text{aug}}, \mathbf{Z}) \in \mathbb{R}^{N_a\times k}. \tag{2}$$

The dilated GCN aggregates messages at multiple dilation rates to capture longer-range topological dependencies while mitigating over-smoothing on dense FC graphs. Together, the Resampler and dilated GCN form an atlas-agnostic graph encoder that provides stable ROI embeddings for subsequent coarse-to-fine harmonization.

Prompt-guided coarse-to-fine harmonization. To avoid unstable global correction, we perform prompt-guided harmonization in a coarse-to-fine manner, where subnetwork-level supervision anchors ROI-level refinement. This design reflects that acquisition-induced heterogeneity can appear at both functional subnetwork and ROI-to-ROI scales.

A single global correction may treat these scales equally, whereas UniFCHarm first harmonizes subnetwork-level connectivity and then refines residual ROI-level bias under the coarse constraint. Given ROI embeddings $\mathbf{H} \in \mathbb{R}^{N_a \times k}$, we encode the target acquisition setting prompt into tokens $\mathbf{C} \in \mathbb{R}^{L \times k}$ using the frozen text encoder in Stage 1 followed by a lightweight adapter with learnable tokens, where L denotes the prompt token length. We inject $\mathbf{C}$ via cross-attention in the fusion modules. We first perform cross-attention on coarse level subnetworks:

$$\mathbf{H}_{\text{coarse}} = \text{Fusion}_c(\mathbf{H}, \mathbf{C}) \in \mathbb{R}^{N_a \times k}, \tag{3}$$

where $\text{Fusion}_c(\cdot)$ uses ROI features as queries and prompt tokens as keys and values, producing prompt-guided node updates while keeping ROI granularity.

We then build a functional subnetwork-level anchor for coarse supervision. Using a fixed Yeo-7 [25] ROI-to-network assignment mask $\mathbf{M} \in \{0,1\}^{N_c \times N_a}$ ($N_c = 7$), ROI features are averaged within each network to form network embeddings:

$$\mathbf{Z}_{\text{net}} = \bar{\mathbf{M}}\mathbf{H}_{\text{coarse}} \in \mathbb{R}^{N_c \times k}, \qquad \bar{\mathbf{M}}_{i,j} = \frac{\mathbf{M}_{i,j}}{\sum_{t=1}^{N_a} \mathbf{M}_{i,t}}. \tag{4}$$

The row-wise normalization implements mean pooling. Yeo-7 serves as a shared cross-atlas functional scaffold that maps ROIs from different atlases into the same seven subnetworks before ROI-level refinement. From $\mathbf{Z}_{\text{net}}$, a subnetwork-level FC $\mathbf{A}_{\text{coarse}} \in \mathbb{R}^{N_c \times N_c}$ is constructed via pairwise similarity between network embeddings, and is used to compute the coarse loss. The model then refines at ROI granularity conditioned on the same prompt:

$$\mathbf{H}_{\text{fine}} = \text{Fusion}_f(\mathbf{H}_{\text{coarse}}, \mathbf{C}) \in \mathbb{R}^{N_a \times k}, \tag{5}$$

where $\text{Fusion}_f(\cdot)$ applies the same cross-attention principle at ROI-level and is optimized by the fine loss. Finally, the harmonized FC $\widehat{\mathbf{A}}$ is reconstructed from $\mathbf{H}_{\text{fine}}$ via pairwise similarity, yielding a single refined FC output.

2.3 Loss Function

To balance harmonization compliance with identity preservation, we optimize a two-term objective:

$$\mathcal{L} = \lambda \mathcal{L}_{\text{coarse}} + \mathcal{L}_{\text{fine}}. \tag{6}$$

$\mathcal{L}_{\text{coarse}}$ stabilizes coarse-level harmonization by maximizing the Pearson correlation between subnetwork-level FCs of the original input and the coarse output:

$$\mathcal{L}_{\text{coarse}} = 1 - \text{corr}\left(\bar{\mathbf{M}}\mathbf{A}_{\text{orig}}\bar{\mathbf{M}}^{\top},\ \mathbf{A}_{\text{coarse}}\right), \tag{7}$$

where $\mathbf{M}$ denotes the ROI-to-subnetwork assignment mask. The original FC is pooled as $\bar{\mathbf{M}}\mathbf{A}_{\text{orig}}\bar{\mathbf{M}}^{\top} \in \mathbb{R}^{N_c \times N_c}$, which is dimensionally consistent with $\mathbf{A}_{\text{coarse}}$. $\mathcal{L}_{\text{fine}}$ preserves identity during ROI-level refinement by matching ROI embeddings extracted from $\mathbf{A}_{\text{orig}}$ and $\widehat{\mathbf{A}}$ via the same Resampler and dilated GCN (Eq. (2)):

$$\mathcal{L}_{\text{fine}} = \left\|\mathbf{H}_{\text{fine}} - \mathbf{H}_{\text{orig}}\right\|_F^2. \tag{8}$$

3 Experiments and Results

3.1 Datasets and Implementation Details

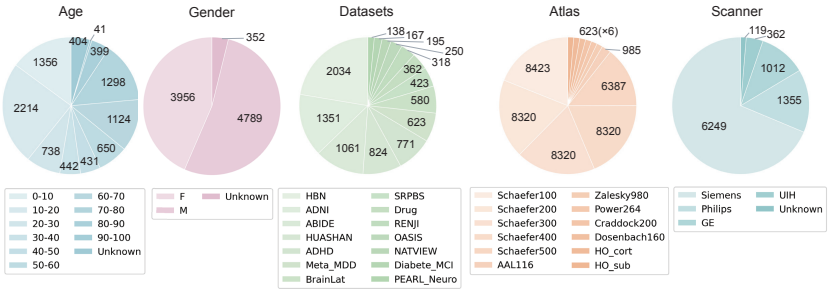

Fig. 2. Dataset overview. Participant distributions are summarized by gender, dataset source, and scanner vendor, with FC counts reported per atlas.

We collect a large-scale, multi-scanner fMRI dataset comprising 44,493 FC-prompt pairs from 9,097 participants, collected from 12 public repositories [2, 4, 5, 6, 7, 9, 14, 15, 17, 18, 19, 24] and 2 internal institutions (Fig. 2). The SRPBS data used for pre-training refer to the SRPBS Multi-disorder MRI Dataset [18]. In contrast, the SRPBS Traveling Subject MRI Dataset [18], which contains 411 scans, is used only for harmonization evaluation and excluded from training. All fMRI scans are preprocessed using a standardized pipeline implemented in Gretna [21], including discarding initial volumes, motion correction, spatial normalization to a common template space, nuisance regression, and temporal filtering. ROI time series are extracted for each atlas, and FC matrices are computed using Pearson correlation.

Experiments are conducted on one NVIDIA A100 GPU. Stage 1 and Stage 2 training take approximately 6.4 and 6.1 days, respectively. In Stage 1, the FC and text encoders are trained for 500 epochs with Adam [11]. In Stage 2, PGH is trained from scratch for 100 epochs on multi-atlas FCs using Adam with batch size 45 and an initial learning rate of 1×10^{-4} halved every 10 epochs; the coarse-loss weight is selected as $\lambda = 0.2$ from $\{0.1, 0.2, \ldots, 1.0\}$. Harmonization is evaluated using four criteria. **Paired fidelity** is quantified by the Pearson correlation (Corr) between the harmonized FC $\widehat{\mathbf{A}}$ and the ground-truth target FC, and by the mean absolute error (MAE) with respect to the same target. **Scanner-effect removal** is measured by 1) the residual scanner significance ratio (Sig-edge), *i.e.*, the percentage of FC edges that remain significantly associated with scanner after harmonization (FDR correction, q=0.05), and 2) the median partial R^2 (Med. part. R^2) across edges, where partial R^2 denotes the fraction of edge-wise variance independently explained by scanner after controlling for covariates. **Target-scanner match** is assessed by training a scanner-condition classifier and reporting the accuracy of predicting the intended per-sample target condition from $\widehat{\mathbf{A}}$. **Biological preservation** is evaluated by age-prediction MAE. For downstream diagnosis, classification accuracy (ACC) and area under the receiver operating characteristic curve (AUC) are reported.

Table 1. Quantitative comparison on SRPBS traveling-subject.

Method	Paired fidelity		Scanner-effect removal ↓		Target-scanner match ↑	Age preservation
	Corr↑	MAE↓	Sig-edge (%)	Med. part. R^2 (%)	Target-scanner ACC (%)↑	Age MAE↓
ComBat [10]	0.762±0.016	0.045±0.003	6.67±1.83	2.54±0.23	33.45±4.52	4.85±0.24
CovBat [3]	0.771±0.014	0.044±0.002	5.85±1.20	2.12±0.15	34.12±5.10	4.72±0.31
ICVAE [16]	0.684±0.028	0.052±0.005	4.12±0.95	1.45±0.12	52.80±8.45	5.25±0.45
SMA [27]	0.725±0.035	0.049±0.004	4.88±1.10	1.82±0.18	48.15±7.22	5.10±0.38
UniFCHarm (Ours)	**0.822±0.011**	**0.033±0.001**	**1.15±0.35**	**0.38±0.05**	**91.25±6.18**	**4.10±0.15**
A1: one-hot embedding	0.785±0.020	0.040±0.003	2.85±0.75	1.05±0.10	72.40±6.50	4.55±0.25
A2: w/o coarse-to-fine constraint	0.791±0.024	0.039±0.004	2.05±0.65	0.85±0.08	84.60±9.12	4.95±0.30
A3: w/o cross -atlas sharing	0.775±0.021	0.041±0.003	3.15±0.82	1.12±0.12	81.45±7.60	4.88±0.28

Note: Sig-edge is computed after FDR correction ($q = 0.05$) across edges.

3.2 Comparison Experiments

We compare UniFCHarm against four representative FC harmonization methods: 1) ComBat [10], 2) CovBat [3], 3) ICVAE [16], and 4) SMA [27]. ComBat and CovBat represent statistical harmonization, ICVAE represents deep latent-variable harmonization, and SMA represents distribution-shift correction via subsampling maximum mean discrepancy matching. Unless specified, we use the default configuration of each comparison method and apply it independently for each atlas. Each baseline is trained or fitted separately for each source-target harmonization task under the same preprocessing and evaluation protocol. This ensures a fair comparison against UniFCHarm, which uses a single unified model across atlases.

Quantitative Comparison. Tab. 1 demonstrates that UniFCHarm consistently outperforms representative harmonization baselines on SRPBS traveling-subject. UniFCHarm attains the highest paired fidelity with MAE 0.033 ± 0.001 while substantially suppressing residual scanner effects, reducing Sig-edge to $1.15 \pm 0.35\%$ and Med. part. R^2 to $0.38 \pm 0.05\%$. Importantly, the target-scanner matching accuracy reaches $91.25 \pm 6.18\%$, validating that semantic prompts provide controllable, acquisition-aware harmonization rather than an unconstrained domain shift. Meanwhile, the lowest age MAE (4.10 ± 0.15) indicates that the hierarchical harmonization avoids over-correcting subject-specific biological variability. The ablations further support our design choices: replacing structured prompts with scanner-specific one-hot embeddings degrades target-scanner matching and increases residual site effects, while dropping the coarse-to-fine constraint or cross-atlas sharing consistently weakens fidelity, scanner-effect removal, and controllability, confirming the benefit of prompt-guided hierarchical harmonization with cross-atlas sharing.

Qualitative Comparison. Fig. 3(A) summarizes qualitative traveling-subject evidence of UniFCHarm, where we visualize FCs using Schaefer100 as an example. Compared with baseline harmonization methods, UniFCHarm yields target-consistent connectivity patterns while keeping the modifications localized and moderate, indicating controlled harmonization that reduces scanner-induced discrepancies without introducing widespread distortions.

Ablation Study. We conduct ablations to isolate the contributions of 1) semantic prompt conditioning for directed harmonization, 2) the coarse-to-fine constraint for stable harmonization, and 3) cross-atlas parameter sharing for unified multi-atlas har-

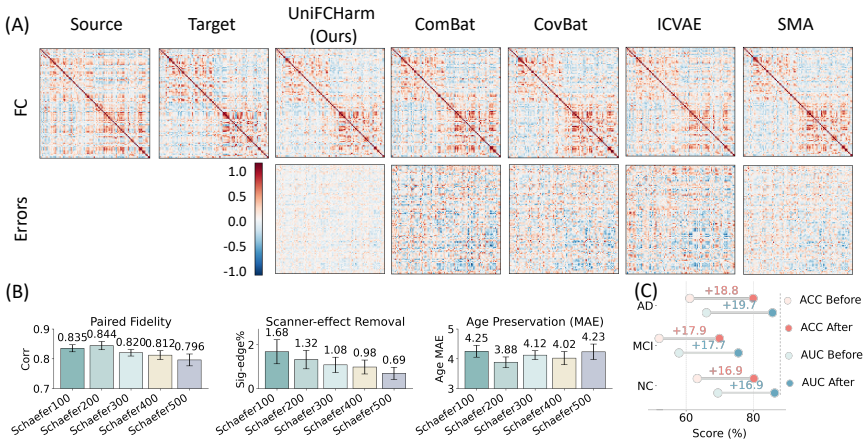

Fig. 3. Performance evaluation of UniFCHarm. (A) Source/target FC, harmonized FC, and residuals (harmonized−target) under specified targets. (B) Multi-atlas generalization (Schaefer100-500) via paired fidelity, scanner-effect removal, and age preservation. (C) Downstream cross-scanner diagnosis improvement (ACC/AUC) for NC, MCI, and AD groups.

monization. Replacing the prompt with a scanner-specific one-hot embedding (A1) drops target-scanner ACC from 91.25% to 72.40% and degrades paired fidelity from Corr/MAE 0.822/0.033 to 0.785/0.040, indicating that directed harmonization requires acquisition-aware prompts rather than coarse scanner labels. Although removing the coarse-to-fine constraint (A2) retains a relatively high target match (84.60%), it weakens scanner-effect removal (Sig-edge 2.05%, Med. part. R^2 0.85%) and harms age preservation (age MAE 4.95), suggesting that coarse-to-fine harmonization stabilizes prompt-guided harmonization and mitigates over-correction of subject-specific signals. Finally, disabling cross-atlas sharing by training atlas-specific PGHs independently (A3) reduces target-scanner match (81.45%) and leaves stronger residual scanner effects (Sig-edge 3.15%) compared with UniFCHarm, supporting the benefit of cross-atlas sharing for consistent harmonization across atlases.

Downstream Analysis. To assess downstream benefits of harmonization under domain shift, we evaluate cross-scanner diagnosis by training on Siemens-derived ADNI FCs (excluded from pretraining in Stage 1) and testing on GE-derived FCs. We use BrainGNN [13] as the downstream classifier. As shown in Fig. 3(C), harmonizing GE-derived FCs to the Siemens domain consistently improves diagnosis across all groups, increasing ACC by 16.9%-18.8% and AUC by 16.9%-19.7%, indicating reduced scanner-induced shifts.

4 Conclusion

In this work, we present UniFCHarm, a unified prompt-guided two-stage framework for multi-scanner and multi-atlas fMRI FC harmonization. UniFCHarm pretrains GraphCLIP to align acquisition- and atlas-aware prompts with FC graphs, yielding prompt

embeddings that guide harmonization toward a user-specified target condition without requiring a single pre-defined target site, while maintaining cross-atlas compatibility. It then performs hierarchical coarse-to-fine harmonization, first adjusting subnetwork-level connectivity and then refining ROI-level edges to remove residual biases, mitigating over-correction while preserving subject-specific FC information. Experiments on traveling-subject benchmark across atlases demonstrate reduced residual site effects and better preserved downstream diagnostic signatures compared with representative baselines. Meanwhile, several limitations remain. First, the current implementation uses Yeo-7 as a fixed shared cross-atlas scaffold, and atlas-matched or learnable scaffolds may further improve ROI-level refinement. Second, future work needs to develop a more mechanistic neuroscientific interpretation of the harmonization.

Acknowledgments. This work was supported in part by National Natural Science Foundation of China (grant numbers 62131015, U23A20295, 82441023, 82394432), the China Ministry of Science and Technology (STI2030-Major Projects-2022ZD0209000, STI2030-Major Projects-2022ZD0213100), The Key R&D Program of Guangdong Province, China (grant number 2023B0303040001), and HPC Platform of ShanghaiTech University.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Alayrac, J.B., Donahue, J., Luc, P., Miech, A., Barr, I., Hasson, Y., Lenc, K., Mensch, A., Millican, K., Reynolds, M., et al.: Flamingo: a visual language model for few-shot learning. Advances in neural information processing systems **35**, 23716–23736 (2022)
2. Alexander, L.M., Escalera, J., Ai, L., Andreotti, C., Febre, K., Mangone, A., Vega-Potler, N., Langer, N., Alexander, A., Kovacs, M., et al.: An open resource for transdiagnostic research in pediatric mental health and learning disorders. Scientific data **4**(1), 1–26 (2017)
3. Chen, A.A., Beer, J.C., Tustison, N.J., Cook, P.A., Shinohara, R.T., Shou, H., Initiative, A.D.N.: Mitigating site effects in covariance for machine learning in neuroimaging data. Human brain mapping **43**(4), 1179–1195 (2022)
4. consortium, A..: The adhd-200 consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. Frontiers in systems neuroscience **6**, 62 (2012)
5. Di Martino, A., Yan, C.G., Li, Q., Denio, E., Castellanos, F.X., Alaerts, K., Anderson, J.S., Assaf, M., Bookheimer, S.Y., Dapretto, M., et al.: The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Molecular psychiatry **19**(6), 659–667 (2014)
6. Ding, D., Zhao, Q., Guo, Q., Liang, X., Luo, J., Yu, L., Zheng, L., Hong, Z., SAS, S.A.S.: Progression and predictors of mild cognitive impairment in chinese elderly: a prospective follow-up in the shanghai aging study. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring **4**, 28–36 (2016)
7. Dzianok, P., Kublik, E.: Pearl-neuro database: Eeg, fmri, health and lifestyle data of middle-aged people at risk of dementia. Scientific Data **11**(1), 276 (2024)
8. Hu, F., Lucas, A., Chen, A.A., Coleman, K., Horng, H., Ng, R.W., Tustison, N.J., Davis, K.A., Shou, H., Li, M., et al.: Deepcombat: A statistically motivated, hyperparameter-robust, deep learning approach to harmonization of neuroimaging data. Human brain mapping **45**(11), e26708 (2024)

9. Jack Jr, C.R., Barnes, J., Bernstein, M.A., Borowski, B.J., Brewer, J., Clegg, S., Dale, A.M., Carmichael, O., Ching, C., DeCarli, C., et al.: Magnetic resonance imaging in alzheimer's disease neuroimaging initiative 2. Alzheimer's & Dementia **11**(7), 740–756 (2015)
10. Johnson, W.E., Li, C., Rabinovic, A.: Adjusting batch effects in microarray expression data using empirical bayes methods. Biostatistics **8**(1), 118–127 (2007)
11. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. International Conference on Learning Representations (ICLR) (2015), https://arxiv.org/abs/1412.6980, arXiv:1412.6980
12. Li, X., Guo, N., Li, Q.: Functional neuroimaging in the new era of big data. Genomics, Proteomics & Bioinformatics **17**(4), 393–401 (2019)
13. Li, X., Zhou, Y., Dvornek, N., Zhang, M., Gao, S., Zhuang, J., Scheinost, D., Staib, L.H., Ventola, P., Duncan, J.S.: Braingnn: Interpretable brain graph neural network for fmri analysis. Medical image analysis **74**, 102233 (2021)
14. Liu, M., Wang, Y., Zhang, H., Yang, Q., Shi, F., Zhou, Y., Shen, D.: Multiscale functional connectome abnormality predicts cognitive outcomes in subcortical ischemic vascular disease. Cerebral Cortex **32**(21), 4641–4656 (2022)
15. Marcus, D.S., Wang, T.H., Parker, J., Csernansky, J.G., Morris, J.C., Buckner, R.L.: Open access series of imaging studies (oasis): cross-sectional mri data in young, middle aged, nondemented, and demented older adults. Journal of cognitive neuroscience **19**(9), 1498–1507 (2007)
16. Moyer, D., Ver Steeg, G., Tax, C.M., Thompson, P.M.: Scanner invariant representations for diffusion mri harmonization. Magnetic resonance in medicine **84**(4), 2174–2189 (2020)
17. Prado, P., Medel, V., Gonzalez-Gomez, R., Sainz-Ballesteros, A., Vidal, V., Santamaría-García, H., Moguilner, S., Mejia, J., Slachevsky, A., Behrens, M.I., et al.: The brainlat project, a multimodal neuroimaging dataset of neurodegeneration from underrepresented backgrounds. Scientific Data **10**(1), 889 (2023)
18. Tanaka, S.C., Yamashita, A., Yahata, N., Itahashi, T., Lisi, G., Yamada, T., Ichikawa, N., Takamura, M., Yoshihara, Y., Kunimatsu, A., et al.: A multi-site, multi-disorder resting-state magnetic resonance image database. Scientific data **8**(1), 227 (2021)
19. Telesford, Q.K., Gonzalez-Moreira, E., Xu, T., Tian, Y., Colcombe, S.J., Cloud, J., Russ, B.E., Falchier, A., Nentwich, M., Madsen, J., et al.: An open-access dataset of naturalistic viewing using simultaneous eeg-fmri. Scientific Data **10**(1), 554 (2023)
20. Tian, D., Zeng, Z., Sun, X., Tong, Q., Li, H., He, H., Gao, J.H., He, Y., Xia, M.: A deep learning-based multisite neuroimage harmonization framework established with a traveling-subject dataset. NeuroImage **257**, 119297 (2022)
21. Wang, J., Wang, X., Xia, M., Liao, X., Evans, A., He, Y.: Gretna: a graph theoretical network analysis toolbox for imaging connectomics. Frontiers in human neuroscience **9**, 386 (2015)
22. Wang, Y.W., Chen, X., Yan, C.G.: Comprehensive evaluation of harmonization on functional brain imaging for multisite data-fusion. NeuroImage **274**, 120089 (2023)
23. Yamashita, A., Yahata, N., Itahashi, T., Lisi, G., Yamada, T., Ichikawa, N., Takamura, M., Yoshihara, Y., Kunimatsu, A., Okada, N., et al.: Harmonization of resting-state functional mri data across multiple imaging sites via the separation of site differences into sampling bias and measurement bias. PLoS biology **17**(4), e3000042 (2019)
24. Yan, C.G., Chen, X., Li, L., Castellanos, F.X., Bai, T.J., Bo, Q.J., Cao, J., Chen, G.M., Chen, N.X., Chen, W., et al.: Reduced default mode network functional connectivity in patients with recurrent major depressive disorder. Proceedings of the National Academy of Sciences **116**(18), 9078–9083 (2019)
25. Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., et al.: The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of neurophysiology (2011)

26. Yu, M., Linn, K.A., Cook, P.A., Phillips, M.L., McInnis, M., Fava, M., Trivedi, M.H., Weissman, M.M., Shinohara, R.T., Sheline, Y.I.: Statistical harmonization corrects site effects in functional connectivity measurements from multi-site fmri data. Human brain mapping **39**(11), 4213–4227 (2018)
27. Zhou, H.H., Singh, V., Johnson, S.C., Wahba, G., Initiative, A.D.N., Berkeley, DeCArli, C.: Statistical tests and identifiability conditions for pooling and analyzing multisite datasets. Proceedings of the National Academy of Sciences **115**(7), 1481–1486 (2018)