CONNECT-4: Brain Connectivity-Guided Hyperedge Graph Fusion for Structural MRI to 4D Rest Functional MRI Synthesis

Salma Hassan[1] ✉[0009−0006−5498−8319], Mostafa Salem[1,2][0000−0001−9915−8390], Ayman Elsayed[3][0000−0003−3381−066X], Olivier Oullier[1,4,5][0009−0007−6271−5019], and Mohammad Yaqub[1][0000−0001−6896−1105]

[1] Division of Computing and Mathematical Sciences, Mohamed Bin Zayed University of Artificial Intelligence, Abu Dhabi, UAE
{Salma.Hassan, Mostafa.Salem, Olivier.Oullier, Mohammad.Yaqub}@mbzuai.ac.ae
[2] Department of Computer Science, Faculty of Computers and Information, Assiut University, Assiut, Egypt
[3] Sheikh Shakhbout Medical City, Abu Dhabi, UAE
ayaahmed@ssmc.ae
[4] Inclusive Brains, Marseille, France
[5] Institute for Artificial Intelligence, Biotech Dental Group, Salon de Provence, France

Abstract. Resting-state functional MRI (rs-fMRI) provides sensitive network-level biomarkers for a broad range of brain disorders, yet its cost, duration, and limited availability mean that most patients undergo only structural MRI (sMRI), leaving a structural–functional gap in routine practice. We propose CONNECT-4, a hyperedge-based multimodal fusion framework that generates subject-specific 4D rs-fMRI volumes directly from T1-weighted sMRI. CONNECT-4 provides three main contributions: (1) a 4D structural to functional generation pipeline that synthesizes full spatiotemporal BOLD sequences from a single structural scan; (2) a hyperedge multimodal fusion module that constructs a unified hypergraph over anatomy, ROI (region of interest) masks, and connectivity priors to generate temporally coherent BOLD dynamics; and (3) a biologically informed multi-scale loss that jointly aligns voxel-wise signals, ROI-level functional connectivity, and global network properties. CONNECT-4 improves MSE by 57.1%, SSIM by 14.6%, and ROI correlation by 17.8% compared to baselines. When synthetic rs-fMRI is added to sMRI in downstream clinical tasks for dementia classification (5 datasets, 12,653 scans) and epilepsy detection, F1 scores improve by 3.3% and 2.2%, respectively, compared to sMRI-only training. These results suggest hyperedge-based multimodal fusion bridges routine imaging and functional diagnostics. Code available at GitHub.

Keywords: fMRI · Graph Neural Networks · Diffusion · Neuroimaging

1 Introduction

Resting-state functional MRI (rs-fMRI) stands at the forefront of modern neuroimaging, offering a window into brain network disruptions 5–10 years before physical atrophy becomes visible [12]. In an era of emerging disease-modifying therapies, this temporal window represents a critical opportunity where early intervention could fundamentally alter disease trajectories before irreversible neurodegeneration occurs. However, a significant structural–functional gap persists in clinical practice due to substantial practical barriers: the extended acquisition time as protocols target at least 10–15 minutes of continuous fMRI data on top of structural imaging, and higher costs ($1,500–$3,000 for rs-fMRI versus $500–$800 for sMRI) [13,15]. These constraints, along with limited availability in most centers, mean that the majority of patients only receive a structural scan. This disparity is evident even in the largest research cohorts like the Alzheimer's Disease Neuroimaging Initiative (ADNI), where only 20% of cases include functional data [18]. As a result, clinicians often make prognostic and diagnostic decisions without functional biomarkers that could improve accuracy [9].

Bridging this gap requires moving beyond traditional medical image synthesis. While existing methods excel at sequence-to-sequence translation, such as mapping T1 to T2 scans by treating the problem as a voxel-wise contrast change [16], fMRI synthesis is a more complex systems-level challenge. Recent efforts have attempted to predict task-evoked activations from rs-fMRI [20] or infer connectivity matrix from structural connectomics [27], but generating full spatiotemporal volumes remains a major hurdle. Early attempts using CycleGANs were limited to treating fMRI as a 3D volume, ignoring the temporal dimension and lacking the quantitative rigor needed for clinical utility [1].

The feasibility of this predictive task is rooted in a robust neurobiological foundation: resting-state functional connectivity is partially constrained by underlying structural connectivity [8,11]. True rs-fMRI reconstruction is inherently challenging, as dynamic physiological states are only partially constrained by static neuroanatomy, necessitating rich structural priors and learned spatiotemporal dynamics. Factors such as vigilance, arousal drifts, and respiration, alongside neurovascular coupling and scanner differences, can systematically shift rs-fMRI measures independently of the underlying structure [25]. While recent deep learning advancements, including generative Graph Convolutional Networks, have predicted functional networks from radiomics-based morphological data, these models often rely on pooling mechanisms that can lose intricate topological connectivity [22]. To our knowledge, this is the first work to move beyond connectivity-matrix prediction to generate full 4D rs-fMRI volumes from sMRI, leveraging anatomical structure as a scaffold that constrains functional interactions and preserving temporally coherent BOLD dynamics through multimodal fusion of structural priors, and can be extended to task-based fMRI paradigms.

To address these challenges, we introduce a framework that formulates rs-fMRI synthesis as a systems-level prediction problem, learning subject-specific spatiotemporal patterns conditioned directly on anatomy. The main contributions of this work are:

1. **4D Structural-to-Functional Synthesis:** We design the first model capable of generating subject-specific 4D rs-fMRI volumes from a single sMRI scan, demonstrating robust generalization across multi-cohort settings.
2. **Hyperedge Multimodal Fusion:** We introduce a novel fusion module that integrates sMRI, ROI masks, and connectivity networks into a unified hypergraph representation to ensure spatiotemporal coherence.
3. **Biologically Informed Loss Design:** We propose a composite loss that jointly aligns voxel signals, ROI connectivity, and global network properties to guide the generative process.

2 Methodology

Processing Steps. Rs-fMRI were co-registered to the processed T1-weighted (T1w) images, followed by slice timing correction, motion correction, spatial smoothing, and temporal filtering, using [7]. Then, we process the T1w sMRI

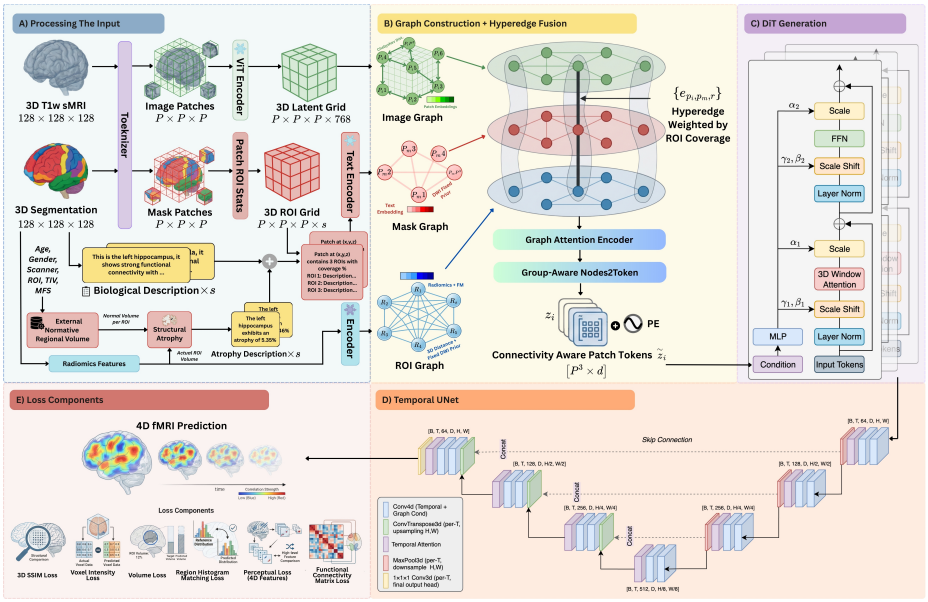

Fig. 1. CONNECT-4 methodology. (A) T1w MRI and its segmentation mask are patchified: image patches are embedded with a ViT encoder; mask patches are turned into biologically guided textual descriptions (structural distribution, coverage, center of mass, and normative context), and ROI-wise radiomics and foundational model features are extracted. (B) Image, mask, and ROI graphs are built; connectivity-guided hyperedges link each patch node to its present ROI nodes with weights proportional to their patch-wise structural distributions, and a graph attention encoder with group-aware Nodes2Token fuses these into patch tokens. (C) These connectivity-aware tokens condition DiT diffusion blocks to synthesize a 4D latent rs-fMRI volume. (D) A temporal graph UNet architecture produces the final 4D rs-fMRI. (E) Loss components.

and its segmentation mask, obtained from a pretrained model [3], by tokenizing both into aligned 3D patches (Fig. 1A). For the image stream, each T1w patch is encoded by a frozen ViT-based brain MRI foundation model (FM) [26]. For the mask stream, each patch is summarized by the distribution of s ROIs it contains. For every ROI present, we derive a structured textual description that combines (i) known functional connectivity of the structure and (ii) normative volume statistics conditioned on age, sex, scanner, and field strength, highlighting deviations relative to the subject's measured volume, based on [19]. Patch-level descriptions aggregate the counts and distributions of ROIs, the patch center of mass, and per-ROI biological summaries; these texts are then embedded using a pretrained clinical text encoder [23]. Finally, we extract radiomics features per ROI and ROI-level features from a FM [2].

Graph Construction and Hyperedges. Using the extracted features, we construct three complementary graphs (Fig. 1B). The image graph is a k-NN graph whose nodes correspond to T1 patches; edges connect the k nearest neighbours under Chebyshev distance between patch centers and are weighted by this distance, while node features are the ViT patch embeddings. The mask graph has the same patch nodes, but edge weights are determined by a fixed diffusion-weighted imaging (DWI) coefficient derived from structural connectivity prior knowledge from external datasets [21]; node features are the text embeddings of the patch-level descriptions. The ROI graph has nodes corresponding to anatomical regions, with features given by the concatenation of T1-based radiomics and FM embeddings, and edges directly taken from the DWI connectivity matrix. Incorporating DWI-derived connectivity is crucial, as it injects white-matter constraints that explicitly couple local anatomy to large-scale network organization, grounding the synthesized rs-fMRI dynamics in biologically plausible pathways rather than purely image-based correlations.

We add a hypergraph that links the graph streams. For each patch p and each ROI r with non-zero coverage in that patch, we create a separate hyperedge $\{e_{p_i,p_m,r}\}$ connecting the T1 patch node p_i, the corresponding mask patch node p_m, and the ROI node r; the hyperedge weight is the fractional coverage of r in the patch, preserving its relative contribution. A shared multi-head graph attention encoder is then applied, where each node updates its representation by aggregating messages from (i) neighbors within its own graph and (ii) nodes connected via incident hyperedges, enabling information to flow both within and across the three graph types. Next, we apply group-aware Nodes2Token tokenization for each patch p, we consider all hyperedges $\{e_{p_i,p_m,r}\}$. Then, patch tokens are defined as $\tilde{z}_i = \sum_r \alpha_r h_r$, where $\alpha_r = \exp(s_r)/\sum_{r'} \exp(s_{r'})$ and s_r is the learned attention score over hyperedge representations h_r, so that ROIs with higher learned importance contribute more strongly. A 3D positional encoding is added to each $\tilde{z}_i$ to yield the final tokens.

Diffusion Transformer Blocks. Our denoising network is parameterized by Diffusion Transformer (DiT) blocks based on DDIM (Fig. 1C). At each diffu-

sion step t, the current noisy latent representation is modeled as a sequence of graph tokens. Each DiT block then applies standard Transformer operations, such as multi-head self-attention, cross-attention, and feed-forward layers, all of which are modulated by the diffusion timestep and conditioning information. Concretely, each block consists of: (i) timestep conditioning, where a sinusoidal or learned embedding of t is added to the token features to make the denoiser time-aware; (ii) self-attention, which enables global spatial and temporal interactions between tokens; (iii) cross-attention to T1-derived and graph/ROI tokens, which aligns the generated rs-fMRI with subject-specific anatomy and regional structure; and (iv) a position-wise MLP that refines each token independently. Stacking multiple DiT blocks yields a deep denoiser that, over the diffusion trajectory, progressively transforms noise into anatomically consistent, temporally coherent rs-fMRI volumes conditioned on the input T1 and graph information.

Temporal UNet. Our time-conditioned UNet transforms DiT outputs into full 4D rs-fMRI volumes (Fig. 1D). DiT-updated graph tokens are first projected onto a spatiotemporal grid, producing feature maps that are fed into a UNet with skip connections. At each resolution, temporal–channel attention blocks apply multi-head attention across time and channels, enabling each voxel prediction to leverage information from all temporal frames to produce temporally coherent, structured rs-fMRI sequences rather than independent frame predictions.

Biological Loss Function Design. Training employs a composite biologically-guided loss that jointly enforces complementary aspects of functional signal fidelity (Fig. 1E): voxel-wise accuracy, structural fidelity (SSIM), temporal coherence, and perceptual realism (4D features from a pretrained fMRI model [5]). Additional losses constrain predictions to brain masks, match regional intensity distributions within ROIs, and preserve the functional connectivity matrix.

3 Experimental Setup

Dataset. We train on two cohorts. Due to limited data availability in public datasets, only T1w sMRI and rs-fMRI are used. First, we utilize the Anti-Amyloid Treatment in Asymptomatic Alzheimer's Disease (A4) study [24], which contains 6,290 paired rs-fMRI and T1w scans spanning the preclinical to full-onset AD. Second, we include a subset of 700 ADNI [18] paired scans. For the first downstream testing task of dementia classification, we evaluate on 5 large public datasets (12,653 scans): A4, ADNI, ANMerge [4], OASIS-3 [14], OASIS-4. For the second task of epilepsy detection, we use the Imaging Database for Epilepsy and Surgery (IDEAS) cohort [28] comprising 542 T1w scans.

Implementation Details. We adopt a standard held-out evaluation protocol with a fixed patient-wise train/validation split on the combined A4+ADNI cohorts and a disjoint test set used exclusively to evaluate synthesis quality. Implementation is based on PyTorch and PyTorch Geometric. All rs-fMRI targets

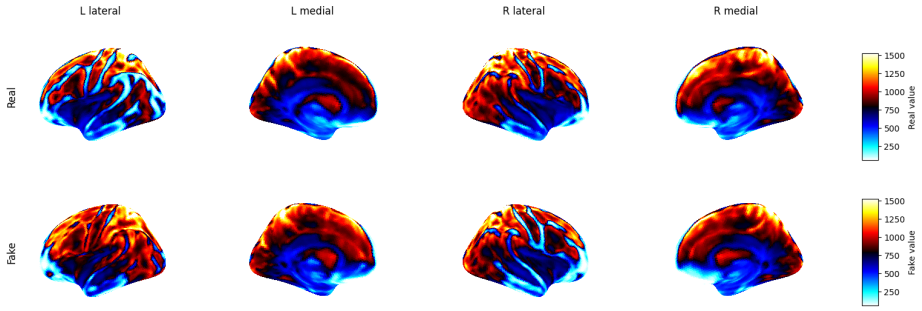

Fig. 2. Qualitative mean 3D rs-fMRI synthesis comparing a subject's real (top row) and CONNECT-4-generated (bottom row) rs-fMRI signal maps, shown for left/right hemispheres in lateral and medial views; color bars indicate the signal intensity scale.

are harmonized to fixed 128-frame sequences (TR = 3s, 3mm isotropic resolution). Models are trained for up to 200 epochs using AdamW, with a batch size of 4. Loss weights were set to $\lambda_{voxel} = 1.0, \lambda_{SSIM} = 0.5, \lambda_{FC} = 0.3, \lambda_{temporal} = 0.2, \lambda_{perceptual} = 0.1$. For rs-fMRI synthesis, we report MSE, SSIM, voxel, ROI and frame-to-frame Pearson correlation, PSNR, and FID/IS. For downstream tasks, we evaluate on completely unseen cohorts using 5-fold cross-validation.

Ablation Studies. We conduct a series of ablation experiments to disentangle the contributions of the different components. First, we compare our approach against several architectural baselines for structural-to-functional synthesis, including Pix2Pix, U-Net2Pix, temporal U-NetFlow, 4D Stable Diffusion, 4D VQ-GAN, and a pure 4D DiT model without any graph conditioning. Second, we study the impact of our graph design by training variants with and without hyperedge fusion, removing individual graph streams, and replacing the hypergraph with simple pairwise graph connections. Third, we evaluate different encoder backbones and graph layers in the graph encoder. Finally, we compare our group-aware Nodes2Token tokenization against simpler pooling schemes to assess the benefit of attention-based, patch-wise token formation.

Table 1. Quantitative comparison of models across MSE, voxel, ROI and frame-to-frame (F2F) correlation, SSIM, PSNR, and IS/FID based on 4D features from [29].

Model	MSE ↓	Voxel Corr ↑	ROI Corr ↑	F2F Corr ↑	SSIM ↑	PSNR ↑	IS ↑	FID ↓
Pix2Pix [10]	0.075	0.521	0.598	0.581	0.612	19.745	2.847	28.54
U-Net2Pix [16]	0.072	0.562	0.641	0.612	0.719	19.982	3.125	24.68
Temporal U-NetFlow [6]	0.064	0.598	0.671	0.703	0.621	20.484	3.241	21.92
Temporal U-NetFlow-Noise [6]	0.066	0.589	0.664	0.694	0.658	20.422	3.187	22.41
4D Stable Diffusion [31]	0.098	0.476	0.531	0.558	0.542	18.923	2.654	35.27
4D VQ-GAN [30]	0.097	0.543	0.592	0.623	0.581	19.213	2.851	28.91
4D DiT [17]	0.063	0.568	0.648	0.718	0.637	20.156	3.156	23.85
Ours (No HyperGraph + Loss)	*0.038*	*0.641*	*0.728*	*0.762*	*0.751*	*21.247*	*3.542*	*16.74*
Ours (Full)	**0.027**	**0.787**	**0.791**	**0.783**	**0.824**	**22.145**	**3.891**	**12.18**

Table 2. Stepwise ablation isolating each contribution, all using the same setup.

Variant	MSE ↓	Voxel Corr ↑	ROI Corr ↑	F2F Corr ↑	SSIM ↑	FID ↓
4D DiT (no graph)	0.063	0.568	0.648	0.718	0.637	23.85
+ Image graph only	0.051	0.601	0.672	0.731	0.671	20.93
+ Mask graph (no DWI prior)	0.047	0.618	0.694	0.744	0.698	19.41
+ DWI prior	0.043	0.629	0.711	0.751	0.724	18.12
+ Hyperedge fusion (no bio. loss)	0.038	0.641	0.728	0.762	0.751	16.74
Full CONNECT-4	**0.027**	**0.787**	**0.791**	**0.783**	**0.824**	**12.18**

Table 3. Summary of downstream tasks, modalities, and performance. • indicates modality used; ∘ indicates not used. A GAT GNN classifier was used for all tasks. † : $p < 0.05$ vs. T1w-only (Wilcoxon, Bonferroni-corrected).

Task	# Classes	Modality		Performance			
		T1w	Synthetic rs-fMRI	Accuracy	F1	Precision	Recall
Dementia	5	•	∘	*0.712 ± 0.028*	0.706 ± 0.031	*0.716 ± 0.034*	0.697 ± 0.033
		∘	•	0.704 ± 0.030	*0.712 ± 0.033*	0.706 ± 0.035	*0.718 ± 0.034*
		•	•	**0.746 ± 0.024**†	**0.739 ± 0.027**†	**0.748 ± 0.029**†	**0.731 ± 0.028**†
Epilepsy	2	•	∘	0.768 ± 0.023	0.846 ± 0.019	0.921 ± 0.019	0.783 ± 0.038
		∘	•	*0.777 ± 0.031*	*0.853 ± 0.024*	*0.925 ± 0.022*	*0.791 ± 0.040*
		•	•	**0.786 ± 0.022**†	**0.864 ± 0.018**†	**0.932 ± 0.017**†	**0.805 ± 0.033**†

4 Results and Discussion

Biological Validity. Figure 2 shows strong spatial correspondence between real and synthetic rs-fMRI across hemispheres. The model preserves expected functional organization in primary sensory and motor cortices and default mode network regions. The ROI correlation of 0.791 validates the hyperedge mechanism's ability to encode region-specific functional characteristics.

Synthesis Quality. Table 1 demonstrates CONNECT-4's superior performance. The model achieves 0.027 MSE, 0.824 SSIM, and 0.787 voxel correlation. Compared to the strongest baseline (DiT), CONNECT-4 reduces MSE by 57.1% and improves voxel correlation by 56.2%. The hypergraph architecture with custom loss achieves marked improvements over the non-hypergraph variant, with voxel correlation increasing from 0.641 to 0.787, validating the hyperedge fusion mechanism's ability to integrate anatomical, functional, and connectivity priors. Traditional methods achieve only 0.521-0.562 voxel correlation, highlighting the domain gap between natural image synthesis and brain functional dynamics.

Disease-Specific Functional Signatures. Table 3 reveals differential performance patterns between dementia and epilepsy that showcase how CONNECT-4 captures distinct pathophysiological mechanisms. For dementia, combining T1w with synthetic rs-fMRI achieves 0.746 accuracy and 0.739 F1, representing 3.4% and 3.3% improvements over T1w-only models. Critically, synthetic rs-fMRI alone achieves 0.704 accuracy, nearly matching T1w performance, demonstrating that CONNECT-4 approximates the statistical structure of distributed network

Table 4. Clinical evaluation results for Dementia classification on scans with both T1w and real fMRI available. Grey rows show the gain of adding synthetic fMRI to T1w. † : $p < 0.05$ vs. T1w-only (Wilcoxon, Bonferroni-corrected).

Input Setting	# Classes	Accuracy	Balanced Acc.	F1	Precision	Recall
T1w only	4 Classes: CTL, MCI, Preclinical, AD	0.709 ± 0.021	0.648 ± 0.027	0.652 ± 0.029	0.663 ± 0.031	0.641 ± 0.025
Real rs-fMRI only		0.701 ± 0.022	0.664 ± 0.028	0.666 ± 0.030	0.676 ± 0.032	0.656 ± 0.026
Synthetic rs-fMRI only		0.676 ± 0.024	0.647 ± 0.034	0.646 ± 0.035	0.660 ± 0.033	0.633 ± 0.023
T1w + Real rs-fMRI		**0.752 ± 0.019**†	**0.701 ± 0.036**†	**0.708 ± 0.039**†	**0.725 ± 0.037**†	**0.692 ± 0.038**†
T1w + Synthetic rs-fMRI		*0.739 ± 0.020*†	*0.684 ± 0.040*†	*0.692 ± 0.043*†	*0.713 ± 0.041*†	*0.671 ± 0.042*†
Early-stage evaluation (CTL / Preclinical / MCI only; AD excluded)						
T1w only	3 Classes: CTL, Preclinical, MCI	0.681 ± 0.017	0.639 ± 0.044	0.647 ± 0.018	0.658 ± 0.016	0.636 ± 0.015
Real rs-fMRI only		**0.722 ± 0.045**†	**0.689 ± 0.046**†	**0.699 ± 0.047**†	**0.712 ± 0.048**†	**0.687 ± 0.049**†
Synthetic rs-fMRI only		0.705 ± 0.050	0.671 ± 0.051	0.680 ± 0.052	0.696 ± 0.053	0.666 ± 0.054
T1w + Real rs-fMRI		*0.718 ± 0.055*	*0.683 ± 0.056*	*0.691 ± 0.057*	*0.706 ± 0.058*	*0.677 ± 0.059*
T1w + Synthetic rs-fMRI		0.710 ± 0.060	0.675 ± 0.061	0.684 ± 0.062	0.702 ± 0.063	0.669 ± 0.064

Table 5. Dementia classification performance when training and testing on real vs. synthetic rs-fMRI. Results are reported for both the 4 and 3 class settings.

Train / Test Setting	# Classes	Accuracy	Balanced Acc.	F1	Precision	Recall
Real rs-fMRI → Real rs-fMRI	4 Classes: CTL, MCI, Preclinical, AD	0.701 ± 0.022	0.664 ± 0.028	0.666 ± 0.030	0.676 ± 0.032	0.656 ± 0.026
Real rs-fMRI → Synthetic rs-fMRI		0.668 ± 0.029	0.638 ± 0.035	0.640 ± 0.037	0.651 ± 0.033	0.630 ± 0.041
Synthetic rs-fMRI → Real rs-fMRI		0.665 ± 0.031	0.635 ± 0.036	0.637 ± 0.039	0.648 ± 0.035	0.627 ± 0.043
Synthetic rs-fMRI → Synthetic rs-fMRI		0.676 ± 0.024	0.647 ± 0.034	0.646 ± 0.035	0.660 ± 0.033	0.633 ± 0.023
Early-stage evaluation (CTL / Preclinical / MCI only; AD excluded)						
Real rs-fMRI → Real rs-fMRI	3 Classes: CTL, Preclinical, MCI	0.722 ± 0.045	0.689 ± 0.046	0.699 ± 0.047	0.712 ± 0.048	0.687 ± 0.049
Real rs-fMRI → Synthetic rs-fMRI		0.697 ± 0.052	0.665 ± 0.053	0.675 ± 0.055	0.690 ± 0.051	0.661 ± 0.059
Synthetic rs-fMRI → Real rs-fMRI		0.695 ± 0.054	0.662 ± 0.056	0.672 ± 0.058	0.687 ± 0.053	0.658 ± 0.062
Synthetic rs-fMRI → Synthetic rs-fMRI		0.705 ± 0.050	0.671 ± 0.051	0.681 ± 0.052	0.696 ± 0.053	0.666 ± 0.054

connectivity. For epilepsy, the combined approach achieves 0.786 accuracy and 0.864 F1, improving by 1.8 points over T1w-only models. However, synthetic rs-fMRI alone still outperforms T1w alone, providing complementary information.

Early Biomarker Detection. Table 4 provides validation for early disease detection. T1w combined with synthetic rs-fMRI (0.739 acc., 0.692 F1) approaches T1w combined with real rs-fMRI (0.752 acc., 0.708 F1). However, the real test is using rs-fMRI for early stage classification, where synthetic rs-fMRI alone substantially outperforms T1w alone and nearly matches real rs-fMRI performance. During the early stages, structural atrophy is minimal, yet functional network disruption is detectable. CONNECT-4 approximates functional signals from structural, suggesting that stable, structurally-constrained connectivity components are partially recoverable from sMRI. The improvements in balanced accuracy and recall indicate reduced false negatives, which are critical for early detection, as missing at-risk patients have significant clinical consequences.

Cross-Domain Generalization and Functional Fidelity. Table 5 addresses a fundamental question about the nature of our synthetic rs-fMRI: does it capture genuine functional biomarkers or merely structural correlates? The cross-domain evaluation reveals that models trained on real rs-fMRI achieve 0.668 accuracy when tested on synthetic rs-fMRI, representing a 3.3% degradation from within-domain performance, while the reverse direction shows a compa-

rable drop. The smaller cross-domain gap in early-stage classification further validates that preclinical functional disruption manifests through stable connectivity changes. These results demonstrate that while our synthetic rs-fMRI may not replicate moment-to-moment physiological noise, it captures anatomically grounded connectivity patterns that are predictive of disease.

5 Conclusion

We presented CONNECT-4, a hyperedge-based multimodal fusion framework that generates subject-specific 4D rs-fMRI from sMRI. CONNECT-4 achieves state-of-the-art synthesis quality and demonstrates clinical utility across dementia and epilepsy tasks. The differential performance across disease contexts reveals that CONNECT-4 yields complementary representations that improve classification in disease-specific contexts. These results indicate that sMRI contains information predictive of stable components of functional connectivity when modeled through connectivity-guided hypergraph fusion and biologically informed losses. Future work includes task-based fMRI synthesis and validation across broader populations.

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

References

1. Abramian, D., Eklund, A.: Generating fmri volumes from t1-weighted volumes using 3d cyclegan. arXiv preprint arXiv:1907.08533 (2019)
2. Barbano, C.A., Brunello, M., Dufumier, B., Grangetto, M.: Anatomical foundation models for brain mris. arXiv preprint arXiv:2408.07079 (2024)
3. Billot, B., Greve, D.N., Puonti, O., Thielscher, A., Van Leemput, K., Fischl, B., Dalca, A.V., Iglesias, J.E., et al.: Synthseg: Segmentation of brain mri scans of any contrast and resolution without retraining. Medical image analysis **86**, 102789 (2023)
4. Birkenbihl, C., Westwood, S., Shi, L., Nevado-Holgado, A., Westman, E., Lovestone, S., Consortium, A., Hofmann-Apitius, M.: Anmerge: a comprehensive and accessible alzheimer's disease patient-level dataset. Journal of Alzheimer's Disease **79**(1), 423–431 (2021)
5. Caro, J.O., Fonseca, A.H.d.O., Averill, C., Rizvi, S.A., Rosati, M., Cross, J.L., Mittal, P., Zappala, E., Levine, D., Dhodapkar, R.M., et al.: Brainlm: A foundation model for brain activity recordings. bioRxiv pp. 2023–09 (2023)
6. Disch, N.A., Kirchhoff, Y., Peretzke, R., Rokuss, M., Roy, S., Ulrich, C., Zimmerer, D., Maier-Hein, K.: Temporal flow matching for learning spatio-temporal trajectories in 4d longitudinal medical imaging. arXiv preprint arXiv:2508.21580 (2025)
7. Esteban, O., Markiewicz, C.J., Blair, R.W., Moodie, C.A., Isik, A.I., Erramuzpe, A., Kent, J.D., Goncalves, M., DuPre, E., Snyder, M., et al.: fmriprep: a robust preprocessing pipeline for functional mri. Nature methods **16**(1), 111–116 (2019)

8. Greicius, M.D., Supekar, K., Menon, V., Dougherty, R.F.: Resting-state functional connectivity reflects structural connectivity in the default mode network. Cerebral cortex **19**(1), 72–78 (2009)
9. Hassan, S., Salem, M., Papineni, V.R.K., Elsayed, A., Yaqub, M.: Magnet-ad: Multitask spatiotemporal gnn for interpretable prediction of pacc and conversion time in preclinical alzheimer. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 348–357. Springer (2025)
10. Henry, J., Natalie, T., Madsen, D.: Pix2pix gan for image-to-image translation. Research Gate Publication pp. 1–5 (2021)
11. Honey, C.J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.P., Meuli, R., Hagmann, P.: Predicting human resting-state functional connectivity from structural connectivity. Proceedings of the National Academy of Sciences **106**(6), 2035–2040 (2009)
12. Khan, T.K.: An algorithm for preclinical diagnosis of alzheimer's disease. Frontiers in neuroscience **12**, 275 (2018)
13. Kumar, V., Lee, J., Liu, H.L., Allen, J., Filippi, C., Holodny, A., Hsu, K., Jain, R., McAndrews, M., Peck, K., et al.: Recommended resting-state fmri acquisition and preprocessing steps for preoperative mapping of language and motor and visual areas in adult and pediatric patients with brain tumors and epilepsy. American Journal of Neuroradiology **45**(2), 139–148 (2024)
14. LaMontagne, P.J., Benzinger, T.L., Morris, J.C., Keefe, S., Hornbeck, R., Xiong, C., Grant, E., Hassenstab, J., Moulder, K., Vlassenko, A.G., et al.: Oasis-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and alzheimer disease. medrxiv pp. 2019–12 (2019)
15. Maity, R., Raja Sankari, V.M., Snekhalatha, U., Velu, S., Alahmadi, T.J., Alhababi, Z.A., Alkahtani, H.K.: Early detection of alzheimer's disease in structural and functional mri. Frontiers in Medicine **11**, 1520878 (2024)
16. Moschetto, A., Puglisi, L., Sargood, A., Dell'Acqua, P., Guarnera, F., Battiato, S., Ravì, D.: Benchmarking gans, diffusion models, and flow matching for t1w-to-t2w mri translation. arXiv preprint arXiv:2507.14575 (2025)
17. Peebles, W., Xie, S.: Scalable diffusion models with transformers. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 4195–4205 (2023)
18. Petersen, R.C., Aisen, P.S., Beckett, L.A., Donohue, M.C., Gamst, A.C., Harvey, D.J., Jack Jr, C., Jagust, W.J., Shaw, L.M., Toga, A.W., et al.: Alzheimer's disease neuroimaging initiative (ADNI) clinical characterization. Neurology **74**(3), 201–209 (2010)
19. Potvin, O., Mouiha, A., Dieumegarde, L., Duchesne, S., s Disease Neuroimaging Initiative, A., et al.: Freesurfer subcortical normative data. Data in brief **9**, 732 (2016)
20. Seo, J., Park, D.K., Yoo, S., Cha, J.: Scalable diffusion transformer for conditional 4d fmri synthesis. arXiv preprint arXiv:2511.22870 (2025)
21. Škoch, A., Rehák Bučková, B., Mareš, J., Tintěra, J., Sanda, P., Jajcay, L., Horáček, J., Španiel, F., Hlinka, J.: Human brain structural connectivity matrices–ready for modelling. Scientific Data **9**(1), 486 (2022)
22. Song, Z., Zhu, C., Jiang, M., Ouyang, M., Zheng, Q.: Predicting functional connectivity network from routinely acquired t1-weighted imaging-based brain network by generative u-gcnet. Neurocomputing **611**, 128709 (2025)
23. Sounack, T., Davis, J., Durieux, B., Chaffin, A., Pollard, T.J., Lehman, E., Johnson, A.E., McDermott, M., Naumann, T., Lindvall, C.: Bioclinical modernbert: A state-of-the-art long-context encoder for biomedical and clinical nlp. arXiv preprint arXiv:2506.10896 (2025)

24. Sperling, R.A., Aisen, P.S.: Anti-amyloid treatment in asymptomatic alzheimer's disease (A4) and longitudinal evaluation of amyloid risk and neurodegeneration (LEARN) study longitudinal results. In: Alzheimer's Association International Conference. ALZ (2023)
25. Tagliazucchi, E., Laufs, H.: Decoding wakefulness levels from typical fmri resting-state data reveals reliable drifts between wakefulness and sleep. Neuron **82**(3), 695–708 (2014)
26. Tak, D., Garomsa, B.A., Chaunzwa, T.L., Zapaishchykova, A., Pardo, J.C.C., Ye, Z., Zielke, J., Ravipati, Y., Vajapeyam, S., Mahootiha, M., et al.: A foundation model for generalized brain mri analysis. medRxiv (2024)
27. Tan, Y.F., Liow, J.L., Tan, P.S., Noman, F., Phan, R.C.W., Ombao, H., Ting, C.M.: Sfc-gan: A generative adversarial network for brain functional and structural connectome translation. arXiv preprint arXiv:2501.07055 (2025)
28. Taylor, P.N., Wang, Y., Simpson, C., Janiukstyte, V., Horsley, J., Leiberg, K., Little, B., Clifford, H., Adler, S., Vos, S.B., et al.: The imaging database for epilepsy and surgery (ideas). Epilepsia **66**(2), 471–481 (2025)
29. Wang, M., Xia, J., Ye, W., Liu, E., Peng, K., Feng, J., Liu, Q., Wen, H.: Slim-brain: A data-and training-efficient foundation model for fmri data analysis. arXiv preprint arXiv:2512.21881 (2025)
30. Zhao, A., Xu, M., Shahin, A.H., Wuyts, W., Jones, M.G., Jacob, J., Alexander, D.C.: 4d vq-gan: synthesising medical scans at any time point for personalised disease progression modelling of idiopathic pulmonary fibrosis. arXiv preprint arXiv:2502.05713 (2025)
31. Zhou, X., Liu, J., Yu, S., Yang, H., Li, C., Tan, T., Wang, S.: A diffusion-driven temporal super-resolution and spatial consistency enhancement framework for 4d mri imaging. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 3–12. Springer (2025)