Functional MRI Time Series Generation via Wavelet-Based Image Transform and Spectral Flow Matching for Brain Disorder Identification

Hwa Hui Tew[1] , Junn Yong Loo[1] , Fang Yu Leong[1],
Julia K. Lau[1], Ding Fan[1], Hernando Ombao[3],
Raphaël C.-W. Phan[1], Chee Pin Tan[2], Chee-Ming Ting[1]

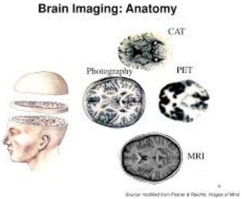
[1]School of Information Technology, Monash University Malaysia
[2]School of Engineering, Monash University Malaysia
[3]Statistics Program, King Abdullah University of Science and Technology

The Fourteenth International Conference on Learning Representations
(ICLR 2026)
Mar 31, 2026

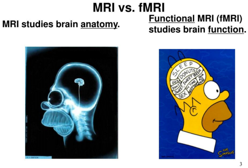
Introduction

Why fMRI signal generation ?

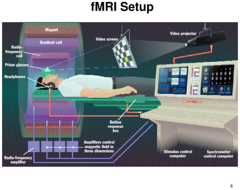
- fMRI measures BOLD dynamics – essential for understanding neural activity in health and disease.
- Clinical fMRI datasets are often small and imbalanced due to acquisition cost and participant burden.
- Existing generative models fail to capture the non-stationary and nonlinear spectral-temporal patterns inherent in BOLD signals.



(a) Brain imaging



(b) MRI vs. fMRI



(c) fMRI setup

Motivation

- fMRI generation
  1. Existing approaches focus on generate brain connectivity directly in the functional connectivity (FC) space.
  2. Such FC representations encode static pairwise relations into dyads and do not effectively capture transient network states within human brain networks
  3. Recent works have revisited time-domain modeling of fMRI as an alternative to correlation-based functional connectivity (FC).
  4. Time-domain approaches still face difficulties in faithfully capturing transient dynamics, multiscale oscillations, and cross-frequency interactions.

How to address these challenges?

- We explore the use of generative techniques to synthesize fMRI signals, enabling data augmentation of available data to enhance brain disorder identification.

Contribution

- We jointly leverage DWT and DCT into a dual-spectral image transform to capture both local and global spatiotemporal-spectral features of fMRI BOLD signals.
- We develop spectral flow matching in the DCT domain to enable efficient coarse-to-fine generation by exploiting the spectral sparsity of fMRI signals.
- Our proposed DSFM demonstrates strong performance on unconditional and conditional spectral image synthesis, and outperforms recent time-series and fMRI generation baselines.

Method: Brain Signal Generation in Frequency Domain

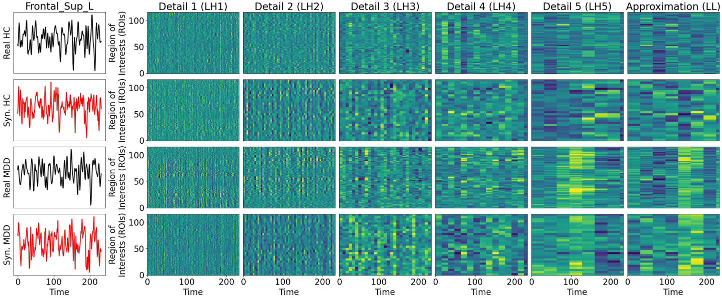


Figure: Ori. vs Syn. BOLD signals and generated normalized scalograms.

- Let $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N] \in \mathbb{R}^{T \times N}$ be multivariate fMRI signals of length T measured from N brain regions
- Time-frequency analysis: (Core Idea) We transform each fMRI signal $\mathbf{y}$ into spectral images, $\mathbf{X}$ via discrete wavelete transfrom (DWT) and discrete cosine transfrom (DCT)
- Generative modeling: To approximate the underlying data distribution $p_{\text{data}}(\mathbf{X})$, and generate synthetic samples $\hat{\mathbf{X}} \sim p_{\boldsymbol{\theta}}(\mathbf{X})$, and then reconstruct time-domain signal $\hat{\mathbf{y}}$ from $\hat{\mathbf{X}}$

Method: Spectral Flow Matching

Forward process progressively adds Gaussian noise to each clean DCT coefficient $z_0[k]$, where $\mathbf{z}_0 \sim p_{\text{data}}$, to generate noisy samples $z_t[k]$ according to *mode-wise conditional perturbation kernel* (Proposition 1.):

$$p(z_t[k] \,|\, z_0[k]) = \mathcal{N}\left(\mu(t,k)\, z_0[k], \sigma(t,k)^2\right), \tag{1}$$

Reparameterization gives

$$z_t[k]|_{z_0[k]} = \mu(t,k)\, z_0[k] + \sigma(t,k)\, \epsilon, \tag{2}$$

- Signal scaling: $\mu(t,k)$
- Noise level: $\sigma(t,k)$

Conditional velocity field differentiating the reparameterized perturbation path from Eq. (2) (Proposition 2):

$$\left.\frac{dz_t[k]}{dt}\right|_{z_0[k]} = v(z_t|z_0;t,k) = \dot{\mu}(t,k)\, z_0[k] + \dot{\sigma}(t,k)\, \epsilon, \tag{3}$$

Spectral flow matching

- **Neural network vector field** v_θ (e.g., transformer) by minimizing simple regression training objective to the conditional velocity field $v(z_t|z_0;t,k)$:

$$\mathcal{L}^{\text{CSFM}}(\theta) = \mathbb{E}_{t, p(z_t \,|\, z_0) p_{\text{data}}(z_0)} \left\| v_\theta(z_t;t,k) - v(z_t|z_0;t,k) \right\|^2. \tag{4}$$

DSFM: fMRI Signal Generation via Flow Matching Model

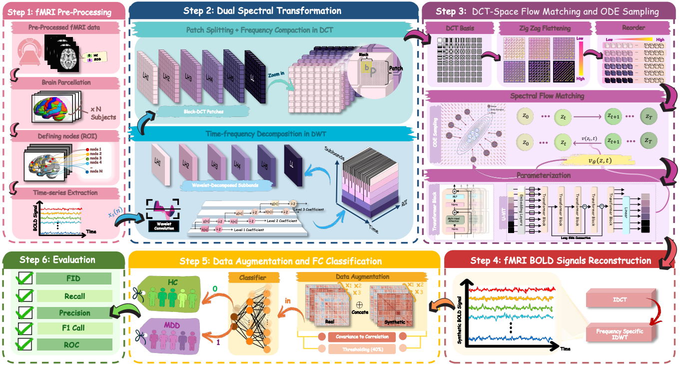


Figure: Flow matching time-series generation based on dual spectral representation

Step 1 **Compute DWT and DCT from fMRI signals**

Step 2 **Train a flow matching model to generate synthetic cosine images**

Step 3 **Reconstruct fMRI signals via iDCT and iDWT**

Step 4 **Functional connectivity (FC) data augmentation with synthetic fMRI signals for brain disorder classification**

Quality of Generated fMRI Signals (Unconditional)

Simulated resting-state fMRI dataset: NetSim (Smith et al., 2011): nodes = 50, TR = 3000 ms, HRF std. dev. = 0.5

Table: Comparison of fMRI signal generation quality with other time-series generative models

	CoT-GAN	DiffTime	DiffWave	TimeVAE	TimeGAN	Diffusion-TS	T2I-Diff	**DSFM**
Context-FID	7.813±.550	0.340±.015	0.244±.018	14.449±.969	0.126±.002	0.105±.006	1.384±.107	0.193±.017
Correlational	26.824±.449	1.501±.048	3.927±.049	17.296±.526	23.502±.039	1.411±.042	4.121±.094	4.552±.041
Discriminative	0.492±.018	0.245±.051	0.402±.029	0.476±.044	0.484±.042	0.167±.023	0.400±.059	0.497±.001
Predictive	0.185±.003	0.100±.000	0.101±.000	0.113±.003	0.126±.002	0.099±.000	0.102±.001	0.104±.000

- Context-Fréchet Inception Distance (Context-FID): Fréchet distance on distributions of time series features from a context-aware time-series encoder
- Correlational score: Absolute error between cross-correlation matrices of real and synthetic time series to assess temporal dependency
- Discriminative score: A classifier can distinguish real from synthetic data
- Predictive score: Mean absolute error of one-step ahead prediction from a post-hoc sequential model trained on synthetic data to predict real data

fMRI Functional Connectivity Classification (Conditional)

We use synthesized resting-state fMRI time series for FC augmentation for brain disorder classification

Dataset: REST-meta-MDD Consortium database (Yan et al., 2019)

- **Participants:** 250 healthy controls (HC) and 227 Major Depressive Disorder (MDD) patients.

Data Collection & Preprocessing:

- Siemens Tim Trio 3T (TR/TE = 2000/30 ms, slice thickness = 3 mm)
- DPARSF pipeline: Time-shifting, despiking, volume registration, alignment to MNI template, smoothing (4mm FWHM), detrending, and temporal bandpass filtering (0.01–0.08 Hz)
- Extract fMRI time series ($T = 232$) of 116 AAL-parcellated ROIs
- Compute spectrogram per ROI based on short-time Fourier Transform
- Compute 116×116 functional connectivity matrix from fMRI time series based on Person correlation

Class-Conditional Generation and FC Augmentation

- DSFM generates synthetic fMRI time series conditioned on class labels (MDD & HC), then used to construct FC matrices.
- **FC data augmentation**: Generate class-specific synthetic FC samples to augment the real FC samples.
- **Classifier Model**: BrainNetCNN, 5-fold stratified cross-validation

Table: The best generation quality and classification performance on the MDD dataset for different generative models using the AAL atlas parcellation, trained on ground-truth data augmented at three levels.

	W/O Aug.	Vanilla-GAN	2D-DCGAN	TimeGAN	Diffusion TS	T2I-Diff	**DSFM (Ours)**
Metric	Real (R.)	R. + Synth. 3×	R. + Synth. 3×	R. + Synth. 1×	R. + Synth. 1×	R. + Synth. 1×	R. + Synth. 1×
Context-FID ↓	–	–	–	4.98 ± 0.65	2.06 ± 0.21	7.45 ± 0.42	1.51 ± 0.41
Correlational ↓	–	–	–	197.05 ± 17.75	64.16 ± 3.92	62.32 ± 1.04	57.30 ± 2.89
Accuracy ↑	58.90 ± 2.98	58.86 ± 2.24	62.88 ± 4.99	66.78 ± 1.66	67.29 ± 1.81	66.87 ± 3.22	70.84 ± 5.89
Recall ↑	58.90 ± 2.98	58.86 ± 2.24	62.88 ± 4.99	66.78 ± 1.66	67.29 ± 1.81	66.87 ± 3.22	70.84 ± 5.89
Precision ↑	59.56 ± 2.74	59.91 ± 2.57	63.12 ± 5.02	67.14 ± 1.70	67.55 ± 1.97	67.06 ± 3.34	70.99 ± 5.80
F1-Score ↑	58.39 ± 3.09	57.64 ± 1.96	62.48 ± 5.25	66.48 ± 1.96	67.21 ± 1.87	66.83 ± 3.21	70.77 ± 5.97
ROC ↑	59.00 ± 2.56	58.57 ± 2.05	62.67 ± 5.15	67.26 ± 2.81	64.57 ± 2.76	67.26 ± 6.00	71.49 ± 5.73

FC Analysis from Synthetic fMRI

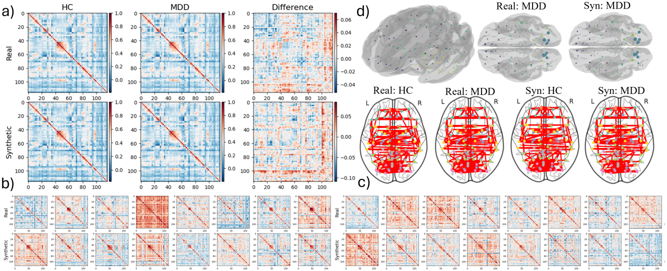


Figure: Real and synthetic connectivity pattern and brain networks.

Table: Similarity between synthetic and real FC networks across FC edges, node strength, and edge betweenness centrality. Higher values indicate better preservation of real FC topology.

Metric	Vanilla-GAN	1D-DCGAN	2D-DCGAN	WGAN	WGAN-GP	**DSFM (Ours)**
FC Edges ↑	0.53 ± 0.06	0.10 ± 0.11	0.54 ± 0.49	0.51 ± 0.47	0.52 ± 0.17	0.99 ± 0.00
Node Strength ↑	0.67 ± 0.08	0.30 ± 0.16	0.53 ± 0.08	0.64 ± 0.09	0.62 ± 0.03	0.99 ± 0.00
Edge Betweenness Centrality ↑	0.11 ± 0.02	0.06 ± 0.02	0.14 ± 0.02	0.14 ± 0.02	0.15 ± 0.02	0.77 ± 0.09

Summary

A framework based on flow matching models for fMRI time-series generation in the frequency domain

- Time-frequency based generative model effectively captures temporal dynamics of frequency content, enabling accurate brain signal generation
- Spectral flow matching improves generation performance over SOTA time-series and fMRI baselines
- Synthetic fMRI signals improve downstream MDD classification, when applied for data augmentation in limited fMRI data settings

Future work:

- Extension to generating time-frequency representation of multivariate signals (in a 3D tensor - time $\times$ frequency $\times$ ROI) with spatio-temporal dependence
- Extension to generating multi-modal neuroimaging signals with different time-frequency resolutions (such as fMRI and EEG)
- Extension to energy-based models to identify out-of-distribution (OOD) patterns in brain spectrograms/scalograms associated with neurological disorders

References

[1] Yuan, Xinyu, and Yan Qiao. “Diffusion-TS: Interpretable diffusion for general time series generation." arXiv preprint arXiv:2403.01742 (2024).

[2] Karras, Tero, et al. “Elucidating the design space of diffusion-based generative models." Advances in neural information processing system (NeurIPS) (2022): 26565-26577.

[3] Song, Yang, et al. “Score-based generative modeling through stochastic differential equations." International Conference on Learning Representations (ICLR), 2021

[4] Kawahara, J. et al. “Brainnetcnn: Convolutional neural networks for brain networks; towards predicting neurodevelopment." NeuroImage 146, 10381049 (2017)

[5] Naiman, I. et al. “Utilizing image transforms and diffusion models for generative modeling of short and long time series. Advances in Neural Information Processing Systems (NeurIPS), 121699â121730 (2024)

[6] Noman, F., Ting, C.M., Kang, H., Phan, R.C.W., Ombao, H. “Graph autoencoders for embedding learning in brain networks and major depressive disorder identification." IEEE Journal of Biomedical and Health Informatics 28(3), 1644â1655 (2024)

Thank You