Structure Discovery in fMRI Time-Series via Dynamic Graph Representation Learning

Luigi Rosati [*1] **Matteo Ciferri** [*1] **Muhammad Kashif** [1] **Dionisia Naddeo** [1] **Nicola Toschi** [1 2]

Abstract

Structured health data often arrive as multivariate time series whose relational structure is unknown, subject-specific, and time-varying. Resting-state fMRI is a salient example: regional BOLD signals provide high-dimensional physiological measurements, but standard functional-connectivity pipelines usually impose a static Pearson-correlation graph. We propose Unsupervised Dynamic Graph Structure Learning (UDGSL), an unsupervised framework that infers a sequence of adjacency matrices directly from windowed fMRI time series. A temporal learner estimates region-region similarities for each window; a GCN encoder and decoder use the learned graph to reconstruct regional signals, yielding dynamic graph representations without labels or predefined connectivity. We evaluate UDGSL on HCP S1200 resting-state fMRI across AAL1 and Brainnetome parcellations. Compared with the best non-UDGSL MLP/LSTM baseline, UDGSL reduces reconstruction MSE by 24-32%; relative to Pearson-correlation graphs, learned graphs reduce forecasting MSE by 26-28% and improve subject-level classification accuracy by 5-7 absolute points. These results suggest that unsupervised graph structure learning can recover useful latent organization in complex physiological time series. We interpret the learned graphs as predictive, data-driven relational representations rather than causal or clinical connectivity estimates, and discuss extensions to signed and static-dynamic graph models.

[*]Equal contribution [1]Department of Biomedicine and Prevention, University of Rome Tor Vergata, Italy [2]A.A. Martinos Center for Biomedical Imaging Harvard Medical School, Massachusetts, USA. Correspondence to: Luigi Rosati <luigi.rosati@uniroma2.eu>.

Proceedings of the Workshop on Structured Data for Health at the 43rd International Conference on Machine Learning, Seoul, South Korea. Copyright 2026 by the author(s).

1. Introduction

Graph-based approaches provide a natural framework for modeling brain systems, where nodes represent regions and edges encode relationships derived from fMRI time-series (Bullmore & Bassett, 2011; Fornito et al., 2016; Smith et al., 2011). This makes fMRI a useful instance of a broader structured-health-data problem: multivariate physiological signals often contain latent, evolving dependencies that are not directly observed. A central challenge in this setting is how to construct the underlying graph. In particular, conventional pipelines rely on predefined measures such as Pearson correlation (Smith et al., 2011), often computed over the entire time-series, resulting in static graphs that fail to capture the inherently time-varying nature of brain interactions (Preti et al., 2017; Allen et al., 2014). Graph Structure Learning (GSL) (Zhu et al., 2021; Kalofolias et al., 2017) offers an alternative by inferring the graph directly from data rather than imposing it a priori (Wu et al., 2022). This is especially relevant in the context of dynamic functional connectivity, where interactions between brain regions evolve over time and cannot be adequately described by a single connectivity matrix. Existing approaches typically rely on sliding-window correlations or biophysical models, but these methods either depend on sensitive hyperparameters or require strong assumptions about the underlying neural dynamics (Hutchison et al., 2013). In this work, we propose an Unsupervised Dynamic Graph Structure Learning (UDGSL) framework to infer time-varying brain graphs directly from fMRI signals from the Human Connectome Project (HCP) S1200 release (Van Essen et al., 2013). Our method leverages a GNN-based encoder–decoder architecture to jointly learn latent representations and dynamic adjacency matrices from preprocessed regional BOLD time-series. By avoiding predefined connectivity measures, the model captures evolving spatial–temporal dependencies in a fully data-driven manner. Compared to traditional correlation-based approaches, our framework improves reconstruction, forecasting, and subject-level classification, highlighting the value of learning relational structure directly from complex structured health time-series rather than imposing a fixed graph a priori.

2. Material and Methods

We propose Unsupervised Dynamic Graph Structure Learning (UDGSL), a framework that learns time-varying graph structures from fMRI time-series while reconstructing the input signal. Given a subject-specific signal X^i, the model is trained by minimizing the reconstruction loss:

$$\mathcal{L}_{\text{rec}}(X^i, \hat{X}^i) = \frac{1}{N_W} \sum_{t=1}^{N_W} \|X^i(t) - \hat{X}^i(t)\|_2^2, \quad (1)$$

where N_W is the number of temporal windows and $\hat{X}^i(t)$ denotes the model reconstruction of the input signal at time window t.

2.1. UDGSL Model

The model consists of three components: a learner f_l, an encoder f_e, and a decoder g. Given a windowed signal $X_W \in \mathbb{R}^{N_W \times W \times V}$, the learner extracts temporal features independently for each brain region using temporal convolutions. The extracted features are then used to compute pairwise similarities, producing a symmetric, non-negative adjacency matrix $A^{(t)}$ for each time window. Node features $F^{(t)}$ are fixed (one-hot encoding of brain regions, see section 2.2), so structural information is primarily captured by the learned adjacency matrices. The model architecture is represented in Figure 1.

The encoder f_e, implemented as a Graph Convolutional Network (GCN), processes the dynamic graphs:

$$H^{(t)} = f_e(A^{(t)}, F^{(t)}), \quad (2)$$

and the decoder g reconstructs the input signal:

$$\hat{X}(t) = g(H^{(t)}). \quad (3)$$

The model is trained end-to-end to reconstruct the original time-series, learning dynamic graph representations without predefined connectivity measures.

The hyperparameter search space and selected values are reported in Appendix B.

2.2. Time-series Reconstruction

The primary training objective is to reconstruct the input signal, making the approach fully unsupervised. The model learns adjacency matrices that encode relationships between regions useful for representing temporal dynamics. No external labels or prior graph structure are required. For this task, node features F are defined as one-hot encoding of brain regions, providing minimal prior information: for each node i and window t, the feature vector $F_i^{(t)}$ is defined as $F_{ij}^{(t)} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$.

2.3. Dynamic Node Prediction

To evaluate the learned graph structure, we perform a forecasting task. For each time step, node embeddings are computed using a GCN:

$$Z^{(t)} = \text{GCN}(A^{(t)}, X^{(t)}), \quad (4)$$

and passed through a recurrent model to capture temporal dependencies:

$$H^{(t)} = \text{RNN}(Z^{(t)}, H^{(t-1)}). \quad (5)$$

Predictions of future node features are obtained via an MLP:

$$\tilde{X}^{(t)} = \text{MLP}(H^{(t)}), \quad (6)$$

and optimized using mean squared error with respect to the next time step.

2.4. Graph Classification

For graph-level tasks, node embeddings are aggregated using average pooling:

$$H_G = \frac{1}{|V|} \sum_{i \in V} H_i, \quad (7)$$

and passed to an MLP for classification:

$$\hat{Y} = \text{MLP}(H_G). \quad (8)$$

The model is trained using cross-entropy loss. In our experiments, this setup is used for subject-level prediction.

3. Results

The primary task used to derive the data-driven adjacency matrices is time-series reconstruction. The resulting learned adjacency sequences are subsequently evaluated in two downstream tasks to assess their representational power. All the experiments are performed on the HCP Dataset. More details about the dataset can be found in Appendix A.

3.1. Time-series Reconstruction

Results for the reconstruction task are shown in Fig. 2 (top) and Table 1.

We compare our approach against standard baselines capturing spatial (MLP), temporal (RNN), and spatio-temporal (RNN+MLP) dependencies. We report parameter counts for all baselines to contextualize model capacity; UDGSL is not larger than the competing MLP/LSTM models.

UDGSL achieved the best performance on both parcellations across all reported reconstruction metrics. These results suggest that explicitly learning dynamic graph structure helps capture spatial and temporal dependencies more effectively than the tested non-graph baselines.

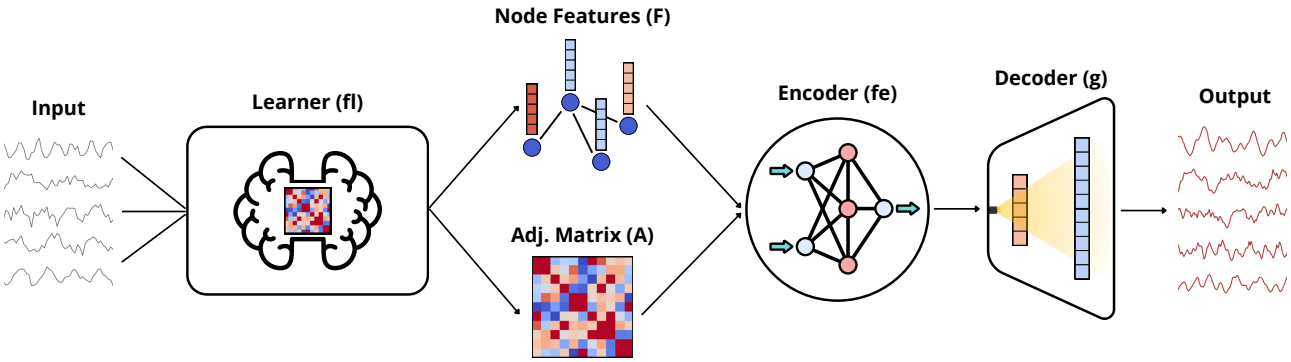

Figure 1. Overview of the UDGSL architecture, highlighting the learner f_l, encoder f_e, and decoder g components. Input time series are first transformed by the learner module f_l, which processes windowed signals producing a graph representation (fixed node features F and learned adjacency matrix A) for each temporal window. These graphs are then fed into the GNN-based encoder f_e. The resulting representations are subsequently passed to the decoder g, which reconstructs a set of time series consistent with the original inputs.

HCP AAL1					
Model	# params	MSE	MAE	SMAPE	PCC
MLP	81268	0.542	0.567	1.139	0.614
LSTM	141484	0.496	0.543	1.071	0.665
LSTM+MLP	102468	0.550	0.571	1.145	0.605
UDGSL	45273	**0.379**	**0.468**	**0.933**	**0.760**
HCP Brainnetome					
Model	# params	MSE	MAE	SMAPE	PCC
MLP	154718	0.586	0.594	1.189	0.572
LSTM	250814	0.552	0.574	1.138	0.605
LSTM+MLP	136918	0.591	0.597	1.192	0.567
UDGSL	51773	**0.376**	**0.467**	**0.931**	**0.762**

Table 1. Time-series reconstruction results. Best values in bold. Results are averaged over 10 seeds; standard deviations are omitted for readability (all ≤ 0.004, i.e., $\lesssim$1% of the mean).

HCP AAL1				
Adj	MSE	PCC	MAE	SMAPE
Pearson	0.677	0.344	0.646	1.213
UDGSL	**0.487**	**0.366**	**0.531**	**1.012**
HCP Brainnetome				
Adj	MSE	PCC	MAE	SMAPE
Pearson	0.674	0.314	0.646	1.210
UDGSL	**0.501**	**0.328**	**0.545**	**1.038**

Table 2. Time-series forecasting results. Best results in bold. Results are averaged over 10 seeds; variability is low ($\leq 4\%$) and omitted.

3.2. Dynamic Node Prediction

We assess the learned adjacency matrices on a node prediction task using the model from Section 2.3. Node features correspond to the original BOLD signals within each time window, and the objective is to predict future normalized BOLD signals from past embeddings. Among the tested RNN variants, the GRU achieved the best performance.

Representative forecasts are shown in Fig. 2 (middle), while quantitative results for both parcellation schemes are reported in Table 2, alongside Pearson-correlation baselines. Averaged over 10 random seeds, UDGSL consistently outperforms the Pearson approach on all metrics.

3.3. Graph Classification

We further evaluate our adjacency matrices on a graph classification task, performing gender classification on the HCP dataset using the models from Section 2.3 and Section 2.4, with one-hot encodings as node features. Among the tested RNN variants, LSTM performs best for Brainnetome, while GRU is optimal for AAL1.

Results for both parcellations, along with comparisons to Pearson-based adjacencies, are reported in Table 3 and Fig. 2 (bottom).

We assess first-order stochastic dominance (FSD) (Hadar & Russell, 1969) using empirical CDFs of test accuracies across runs. For AAL1, the dominance statistic is $S = 0.70$ ($p = 0.009$, 1000 permutations), with a null 95% range of $[0.00, 0.50]$ and a large effect size ($d = 1.58$). For Brainnetome, we obtain $S = 0.65$ ($p = 0.002$), with the same null range and a similarly large effect size ($d = 1.52$). Consistent results across parcellations indicate robustness to variations in the number of nodes.

HCP AAL1				
Adj	Accuracy	F1	Precision	Recall
Pearson	0.72 ± 0.01	0.74 ± 0.02	0.74 ± 0.02	0.75 ± 0.04
UDGSL	$\mathbf{0.79 \pm 0.01}$	$\mathbf{0.80 \pm 0.01}$	$\mathbf{0.83 \pm 0.02}$	$\mathbf{0.78 \pm 0.02}$
HCP Brainnetome				
Adj	Accuracy	F1	Precision	Recall
Pearson	0.69 ± 0.01	0.69 ± 0.01	0.68 ± 0.01	0.76 ± 0.03
UDGSL	$\mathbf{0.74 \pm 0.01}$	$\mathbf{0.74 \pm 0.02}$	$\mathbf{0.78 \pm 0.02}$	$\mathbf{0.78 \pm 0.04}$

Table 3. Gender classification task with different adjacency matrix embeddings. Mean and standard deviations over 10 seeds. Best results for each metric are shown in bold.

Figure 2. (Top row): Time-series reconstruction. For three representative ROIs, we show the reconstructed signals from our model (orange) and the same model using Pearson correlation adjacency matrices (green), compared to the original signal (blue). (Middle row): Node feature prediction. The grey-shaded area represents the forecasting horizon. The models are trained on the preceding time window and tasked with forecasting the brain activity (blue), based on either our learned graph embeddings (orange) or Pearson ones (green). (Bottom row): Gender classification. Left: classification accuracy (± std) across two parcellations (AAL1 and Brainnetome), comparing our learned adjacency matrices (orange) to Pearson-based ones (green). Center and right: Empirical Cumulative Distribution Function (ECDF) plots from the first-order stochastic dominance test, showing that our model significantly outperforms the Pearson baseline in both parcellations.

4. Discussion and Conclusions

We introduced a deep learning framework for learning latent, time-varying structure directly from multivariate fMRI time-series. Unlike approaches based on predefined connectivity or known graphs, our method infers interaction structure from the data, enabling flexible and adaptive representations of complex dynamics.

More broadly, the approach applies to structured health time-series where relational organization is not explicitly available but expected to be informative, such as physiological monitoring, wearable sensing, and electronic health records. In these settings, the model acts as a structure discovery mechanism that jointly learns representations and an adaptive notion of relationships without requiring a predefined graph. By learning structure in an unsupervised manner, the model supports downstream tasks such as prediction and representation learning while remaining agnostic to fixed topologies. This is particularly relevant in healthcare, where uncovering latent dependencies can improve predictive performance and yield interpretable insights into complex physiological systems.

The learned representation provides a compact, structure-aware embedding that can be directly used in downstream tasks. Importantly, the inferred adjacency matrices should be interpreted as predictive relational representations rather than causal estimates or validated clinical biomarkers.

Despite these advantages, the current formulation has limitations. UDGSL is intrinsically dynamic and not directly suited for static connectivity analyses, and the non-negativity constraint prevents explicit modeling of signed interactions such as anticorrelations.

Future work may address these limitations by incorporating joint static–dynamic modeling and signed or mixed-sign adjacency representations, enabling richer relational modeling in healthcare time-series.

Acknowledgements

The authors are grateful to Simeon Spasov and Alexander Campbell for their extensive discussion and valuable insights, which helped shape this work.

References

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., and Calhoun, V. D. Tracking whole-brain connectivity dynamics in the resting state. *Cerebral cortex*, 24(3):663–676, 2014.

Bullmore, E. T. and Bassett, D. S. Brain graphs: graphical models of the human brain connectome. *Annual review of clinical psychology*, 7(1):113–140, 2011.

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A. R., et al. The human brainnetome atlas: a new brain atlas based on connectional architecture. *Cerebral cortex*, 26(8):3508–3526, 2016.

Fornito, A., Zalesky, A., and Bullmore, E. *Fundamentals of brain network analysis*. Academic press, 2016.

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., et al. The minimal preprocessing pipelines for the human connectome project. *Neuroimage*, 80:105–124, 2013.

Hadar, J. and Russell, W. R. Rules for ordering uncertain prospects. *The American economic review*, 59(1):25–34, 1969.

Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., Della Penna, S., Duyn, J. H., Glover, G. H., Gonzalez-Castillo, J., et al. Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage*, 80:360–378, 2013.

Kalofolias, V., Loukas, A., Thanou, D., and Frossard, P. Learning time varying graphs. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2826–2830. Ieee, 2017.

Preti, M. G., Bolton, T. A., and Van De Ville, D. The dynamic functional connectome: State-of-the-art and perspectives. *Neuroimage*, 160:41–54, 2017.

Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., and Smith, S. M. Automatic denoising of functional mri data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage*, 90:449–468, 2014.

Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., Ramsey, J. D., and Woolrich, M. W. Network modelling methods for fmri. *Neuroimage*, 54(2):875–891, 2011.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., and Joliot, M. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage*, 15(1): 273–289, 2002.

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., and Ugurbil, K. The wu-minn human connectome project: An overview. *NeuroImage*, 80:62–79, 2013. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2013.05.041. Mapping the Connectome.

Wu, L., Cui, P., Pei, J., and Zhao, L. *Graph Neural Networks: Foundations, Frontiers, and Applications*. Springer Singapore, Singapore, 2022.

Zhu, Y., Xu, W., Zhang, J., Du, Y., Zhang, J., Liu, Q., Yang, C., and Wu, S. A survey on graph structure learning: Progress and opportunities. *arXiv preprint arXiv:2103.03036*, 2021.

A. HCP Data

We use the Human Connectome Project (HCP) dataset (Van Essen et al., 2013), a large-scale neuroimaging resource that provides high-quality functional brain connectivity data. Specifically, we select subjects from the 1200 Subjects Release (S1200), which includes 3T MRI data from 1206 healthy young adults aged 22–35. Each subject underwent four resting-state fMRI sessions; we selected the first session (LR), containing 1200 time points. We retained subjects who completed the full selected fMRI session, yielding $N = 1093$ subjects. Time-series are processed using the same pipeline as (Glasser et al., 2013). Brain networks are constructed using two different brain parcellations: the Brainnetome Atlas (Fan et al., 2016) and the AAL1 Atlas (Tzourio-Mazoyer et al., 2002), which segment the brain into 246 and 116 regions of interest (ROIs), respectively.
All the data were processed using the HCP minimal preprocessing pipeline, followed by ICA-FIX denoising, as described in detail in (Glasser et al., 2013) and in (Salimi-Khorshidi et al., 2014).

B. Model Hyperparameters

For each candidate configuration, the model is trained on 5 independent dataset splits (each defined by a different random seed). The validation set is used for early stopping, while performance is evaluated on a held-out partition. All candidate configurations are reported in Table 4, and the configuration with the lowest mean MSE across the 5 held-out model selection partitions is chosen (highlighted in bold). The final results are then obtained by retraining the selected configuration on a separate set of 10 independently generated splits using different random seeds.

Hyperparameter	**Explored values**
Preprocessing:	
Window Size	**10**, 20, 30
Window Stride	2, **5**, 10
Learner:	
Hidden Dimension	**20**, 50, 100
Num. Layers	1, 2, 5, 10, **12**, 15
Dropout	**0.0**, 0.1, 0.3
Encoder:	
Hidden Dimension	20, **50**, 100
Num. Layers	1, **2**, 5, 10
Dropout	**0.0**, 0.1, 0.3
GNN family	GAT, **GCN**
Decoder:	
Hidden Dimension	32, **64**, 128
Kernel	2, **4**, 8
Dropout	0.0, 0.1, **0.2**
Training:	
Learning Rate	$1e^{-5}$, $1e^{-4}$, $\mathbf{1e^{-3}}$
Batch Size	4, 8, **16**

Table 4. Hyperparameter search space. Best performing values are shown in bold.

C. Adjacency matrix

To verify that our model is effectively using and building a task-relevant adjacency matrix, we provide alternative matrices to the model to be trained with. By providing an external (non-trained) $A(t)$, we completely skip the Learner module $\boldsymbol{f_l}$, so this is equivalent to performing an ablation study on the core module of our model.

As a baseline for this comparison, we choose correlation coefficients computed as in the correlation-based node features (where negative correlation coefficients have been discarded). In this way $A(t)$ is a matrix with Pearson Correlation

coefficients of the different time-series for the window t. In this case we still keep some useful information in the adjacency, but we also add another baseline in which the model is trained with pure noise on the edges, producing a random $A(t)$ for each window.

Table 5 reports the results of this comparison.

HCP AAL1				
Adj. Matrix	MSE	MAE	SMAPE	PCC
Noise	0.920 ± 0.003	0.760 ± 0.001	1.974 ± 0.004	0.0029 ± 0.0006
Pearson	0.917 ± 0.003	0.759 ± 0.001	1.827 ± 0.004	0.049 ± 0.001
Learner	$\mathbf{0.376 \pm 0.003}$	$\mathbf{0.467 \pm 0.002}$	$\mathbf{0.931 \pm 0.004}$	$\mathbf{0.762 \pm 0.002}$

Table 5. Time-series reconstruction task with different adjacency matrices provided during training. Mean and standard deviations over 10 seeds.

The learned adjacency matrices yield significantly better reconstruction performance. The poor performance in the noise case is expected, as the node features used for the reconstruction task are minimally informative (one-hot encodings of brain regions). Similarly, the low performance of the Pearson adjacency matrix suggests that correlation-based edges alone do not capture sufficient information about the original signal for accurate reconstruction.