

TReND: Transformer derived features and Regularized NMF for neonatal functional network Delineation

Sovesh Mohapatra^{1,2}, Minhui Ouyang^{1,3}, Shufang Tan¹, Jianlin Guo¹,
Lianglong Sun⁴, Yong He⁴, and Hao Huang^{1,3}

¹ Department of Radiology, Children’s Hospital of Philadelphia, PA, United States

² Department of Bioengineering, University of Pennsylvania, PA, United States

³ Department of Radiology, University of Pennsylvania, PA, United States

⁴ State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China
huangh6@chop.edu

Abstract. Precise parcellation of functional networks (FNs) of early developing human brain is the fundamental basis for identifying biomarker of developmental disorders and understanding functional development. Resting-state fMRI (rs-fMRI) enables in vivo exploration of functional changes, but adult FN parcellations cannot be directly applied to the neonates due to incomplete network maturation. No standardized neonatal functional atlas is currently available. To solve this fundamental issue, we propose TReND, a novel and fully automated self-supervised transformer-autoencoder framework that integrates regularized nonnegative matrix factorization (RNMF) to unveil the FNs in neonates. TReND effectively disentangles spatiotemporal features in voxel-wise rs-fMRI data. The framework integrates confidence-adaptive masks into transformer self-attention layers to mitigate noise influence. A self supervised decoder acts as a regulator to refine the encoder’s latent embeddings, which serve as reliable temporal features. For spatial coherence, we incorporate brain surface-based geodesic distances as spatial encodings along with functional connectivity from temporal features. The TReND clustering approach processes these features under sparsity and smoothness constraints, producing robust and biologically plausible parcellations. We extensively validated our TReND framework on three different rs-fMRI datasets: simulated, dHCP and HCP-YA against comparable traditional feature extraction and clustering techniques. Our results demonstrated the superiority of the TReND framework in the delineation of neonate FNs with significantly better spatial contiguity and functional homogeneity. Collectively, we established TReND, a novel and robust framework, for neonatal FN delineation. TReND-derived neonatal FNs could serve as a neonatal functional atlas for perinatal populations in health and disease.

Keywords: Functional Parcellation · Neonate · Transformer

1 Introduction

Neuroscientists have long attempted to subdivide the human brain into a mesh of anatomically and functionally distinct, contiguous regions [1–7]. This challenge become particularly complex in the neonatal brain, where functional organization differs markedly from that of adults. During the third trimester, the neonatal brain undergoes a critical phase of enhanced functional segregation, largely driven by the rapid development of functional connectivity (FC) and the formation of hubs in primary regions [8–10]. This period of growth and organization leads to the emergence of functionally segregated networks, revealing the underlying principles that shape both healthy and diseased brain states [11–13]. However, achieving accurate and reliable parcellation of specific functional networks (FNs) in newborns presents unique challenges. The combined effects of rapid functional segregation, low imaging quality, and the absence of established functional atlases as prior reference knowledge complicate this process [14, 15].

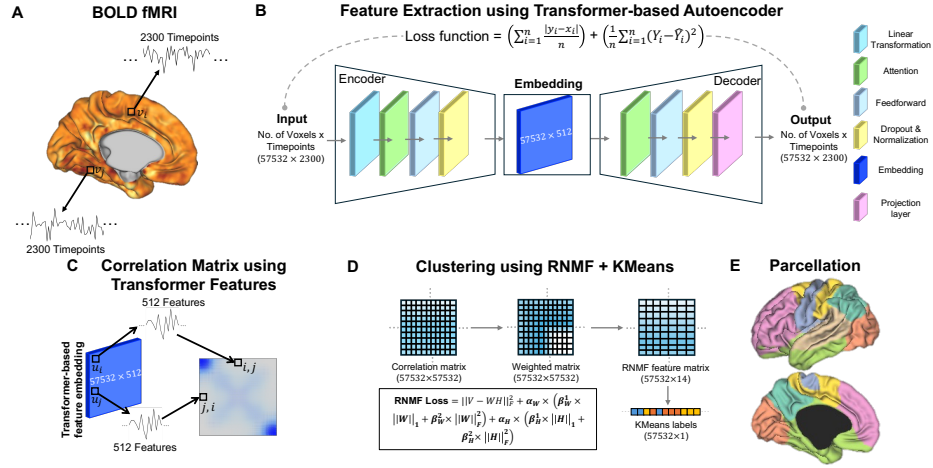


Fig. 1. Schematic representation of end-to-end population-level functional parcellation framework. **A.** Processed Blood Oxygenation Level Dependent (BOLD) fMRI data. **B.** Feature extraction using transformer-based autoencoder to convert BOLD signals into low-dimensional embeddings. **C.** Correlation matrix from the transformer-based feature embeddings. **D.** Clustering using RNMF + KMeans are performed to generate parcellation of different regions. **E.** Population-level parcellation shows different regions based on averaged correlation matrix across subjects.

Recent advances in artificial intelligence have enabled temporal features extraction from fMRI blood oxygenation level-dependent (BOLD) signals using self-supervised and contrastive learning. While supervised deep learning (DL) architectures have learned FC patterns to predict clinical and demographic variables (e.g., [16, 17]), self-supervised architectures such as variational autoen-

coders, which embed FC profiles into low-dimensional latent spaces [18], remain largely unexplored in brain parcellation. Here, we introduce, to our knowledge, the first transformer-based autoencoder that extracts temporal BOLD features for functional parcellation, extending DL-based embeddings beyond prediction to unsupervised parcellation frameworks.

In this study, we developed the Transformer derived Regularized NMF for neonatal FN Delineation (TReND) framework. This approach involves self-supervised DL-based BOLD time course feature extraction with a novel attention mechanism tailored for brain parcellation, improved regularized NMF framework and novel fusion of geodesic spatial constraints with average FC for the delineation of FNs.

2 Methods

2.1 Data description

The study included three datasets. The simulated fMRI dataset was generated using the SimTB toolbox [19], comprising 200 simulated 2D images, each with dimensions of 100x100 voxels and 1600 timepoints, including 15 FNs to replicate real brain connectivity. Rician noise was added to the simulated images with contrast-to-noise ratios ranging from 0.1 to 0.3. The developing human connectome project (dHCP) included 300 term neonates (mean scan age: 41.16 weeks) [20]. All data were preprocessed and registered to the 40 week neonatal surface ($\sim 32K$ vertices per hemisphere, 2300 timepoints) [21]. HCP Young Adult (HCP-YA) comprised of 200 randomly selected subjects and all were preprocessed and registered to fsLR surface with 1200 timepoints [22, 23].

2.2 Overview of TReND framework

TReND is an automated self-supervised transformer-based autoencoder framework that incorporates confidence-adaptive masks to extract salient temporal features and employs geodesic-guided RNMF-KMeans clustering for robust neonatal FN delineation. (Fig 1).

Transformer based autoencoder with confidence-adaptive mask for temporal feature extraction. The architecture is designed to capture long-range temporal dependencies while mitigating the over-influence of noise. The autoencoder consists of two main components: an encoder $f_\theta(\cdot)$ that maps the input fMRI signal (Fig 1A) to a latent representation and a decoder acting as a self-regulator $g_\phi(\cdot)$ that reconstructs the original signal from the latent representation. The encoder is built by stacking multiple transformer layers that capture long-range temporal dependencies via self-attention (Fig 1B and 2A-*left panel*) [24]. The attention operation is being modified to include a confidence-adaptive mask M :

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^\top}{\sqrt{d_k}} + M \right) V \quad (1)$$

Here, Q , K , and V represent the query, key, and value matrices obtained from linear projections of the input d_k (e.g. 512) is the dimensionality of the key vectors, and the confidence mask M modulates the contribution of each token based on its reliability. A central innovation of our approach is the ConfidenceModule:

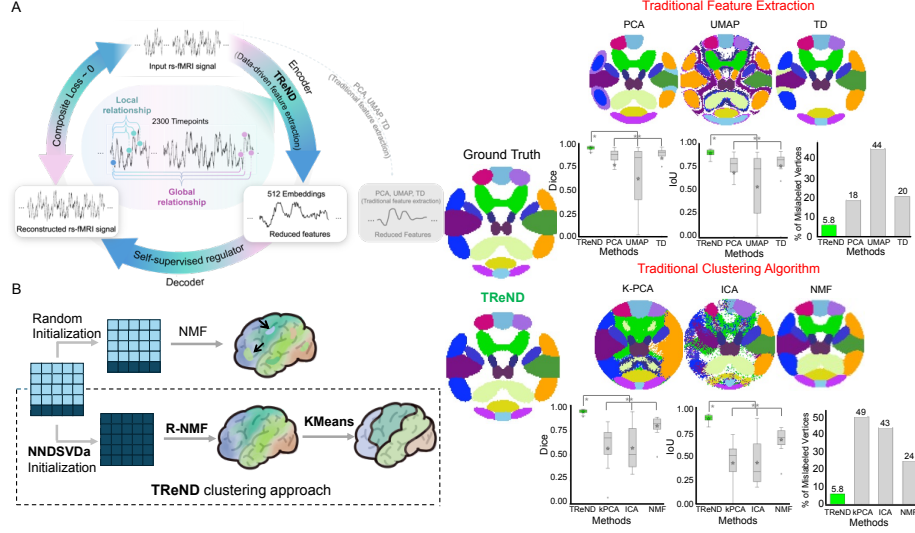


Fig. 2. Performance evaluation of TReND versus traditional parcellation methods on simulated data. **A.** *Left panel:* The cartoon illustrates how TReND extracts features and preserves more fine-grained details. *Right panel:* Comparison of feature extraction techniques: Principal component analysis (PCA), Uniform manifold approximation and projection (UMAP), Tensor decomposition (TD), and TReND. **B.** *Left panel:* The cartoon shows how TReND produces spatially coherent clusters. *Right panel:* Comparison of clustering methods: kernel-PCA (kPCA), Independent component analysis (ICA), NMF, and TReND.

a two-layer multilayer perceptron (MLP) that estimates the reliability of each embedded token. Given an embedded sequence $x_{emb} \in R^{T \times B \times d_k}$ (where B is the batch size, T is the number of time points), the module computes a scalar confidence score for each token. These scores, which range between 0 and 1, are transformed into log domain values and combined pairwise to form the additive confidence mask.

1. Token-wise confidence estimation: Each token is passed through the ConfidenceModule, which applies a ReLU activation followed by a sigmoid function to produce a confidence score that reflects the reliability of the token.
2. Mask formation: The log-transformed confidence scores are summed pairwise to produce a mask \mathbf{M} such that for tokens i and j , the mask entry is given by:

$$M_{ij} = \log(c_i + \epsilon) + \log(c_j + \epsilon) \quad (2)$$

where c_i and c_j denote the confidence scores for tokens i and j respectively, and ϵ is a small constant to avoid numerical instability.

This confidence-adaptive mechanism implicitly detects and down-weights over-influencing tokens during the attention computation, ensuring that the latent representations are dominated by reliable and discriminative features.

The transformer-based autoencoder was trained in a self-supervised manner with a reconstruction objective for 1500 epochs on 4 A100 (80GB) GPUs, using a batch size of 1, Adam optimizer, and learning rate scheduler starting at 1e-4 to extract 512 features. This inherently regularizes the latent space, guiding the encoder to extract the most salient temporal features.

Feature extraction loss. To balance the trade-off between bias and variance in the reconstruction, we employ a composite loss function that combines Mean Squared Error (MSE) and the Mean Absolute Error (MAE):

$$\mathcal{L} = \lambda_{\text{MSE}} \cdot \mathcal{L}_{\text{MSE}} + \lambda_{\text{MAE}} \cdot \mathcal{L}_{\text{MAE}} \quad (3)$$

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2 \quad \mathcal{L}_{\text{MAE}} = \frac{1}{N} \sum_{i=1}^N |x_i - \hat{x}_i| \quad (4)$$

where N is the total number of data points across time and voxels, and λ_{MSE} and λ_{MAE} are hyperparameters that adjust the relative contribution of each loss component.

Regularized non-negative matrix factorization with KMeans for clustering. To robustly parcellate neonatal FNs, we propose a clustering approach that synergistically integrates an advanced initialization strategy with regularization parameters [25]. Our method begins with the construction of the data matrix (\mathbf{V}) which is derived by integrating FC matrix (Fig 1C and 2B-*left panel*) obtained via temporal features with spatial encodings derived from brain surface-based geodesic distances. The \mathbf{V} matrix, constructed by taking the absolute average FC across subjects, is non-negative by design. This constructed data matrix is then decomposed using RNMF.

As NMF algorithms are well known for their sensitivity to initialization; random initialization often leads to suboptimal or unstable solutions. To mitigate this, we employ the non-negative double singular value decomposition with averaging (NNDSVDa) to initialize \mathbf{W} (basis matrix) and \mathbf{H} (encoding matrix) [26]. This approach provides a more reliable starting point than random initialization, helping to position the algorithm in a favorable region of the solution space.

Clustering loss. To extract functionally and spatially coherent regions, we introduce a loss function that both minimizes the reconstruction error and in-

incorporates regularization terms to promote sparsity and smoothness in the decomposed factors:

$$\mathcal{L} = \|V - WH\|_F^2 + \alpha_W \times (\beta_W^1 \times \|W\|_1 + \beta_W^2 \times \|W\|_F^2) + \alpha_H \times (\beta_H^1 \times \|H\|_1 + \beta_H^2 \times \|H\|_F^2) \quad (5)$$

where $\|\cdot\|_F$ denotes the Frobenius norm, ensuring that the overall reconstruction error is minimized. $\|\cdot\|_1$ promotes sparsity in the factors, essential for isolating brain networks. The parameters α_W and α_H control the overall regularization strength for the basis and encoding matrices, respectively. $\beta_{W,H}^1$ and $\beta_{W,H}^2$ balance the respective weighted contributions of the sparsity and smoothness.

After factorizing $V \approx WH$, we obtain a latent representation of each brain voxel in W . To define final parcels, we apply KMeans on W (rows of W correspond to voxels, columns correspond to latent features) (Fig 1D). This additional clustering step refines each row's component assignments into discrete labels, thereby producing coherent, biologically plausible neonatal functional brain parcellations (Fig 1E).

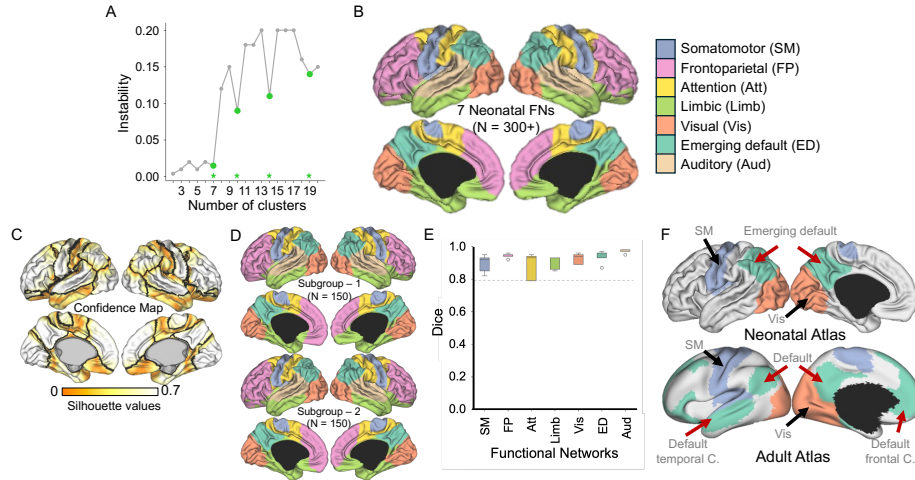


Fig. 3. Coarse parcellation of 7 functional networks in neonates derived from a 300-subject dHCP cohort. **A.** Stability analysis of the clustering algorithm identifies 7 and 19 networks as robust estimates. **B.** Neonatal cortical parcellation of 7 FNs. **C.** Confidence map representing the reliability of the 7 neonatal FNs parcellation. **D.** Discovery and replication of the 7 neonatal FNs cortical parcellation. **E.** Dice score evaluation FN-wise. **F.** Comparison of primary and higher-order FNs in TReND-derived neonate atlas and existing adult atlas [3]. (Abbreviation: C.: Component)

2.3 Evaluation of TReND framework

Given the absence of well-established functional atlases for neonates, we evaluated TReND using a simulated dataset with 15 FNs as ground truth. We then applied to the dHCP cohort with 300 term neonates. To find optimal starting points for delineation of FNs, we assessed cluster stability by splitting the $\sim 64K$ vertices into two groups, clustering them independently, and using learned parameters from one to predict clusters in the other [27]. The agreement between predicted and derived clusters reflects the stability. Subsequently, the spatial confidence map is calculated using the silhouette measure, which quantifies correlation-based similarity to the assigned network versus the nearest alternative [28]. The validation is further performed via bootstrap analysis across 100 combinations using discovery and replication subsets (150 subjects each) of the dHCP cohort. TReND’s applicability is also tested on the HCP-YA dataset using the Yeo 7-Network atlas [3].

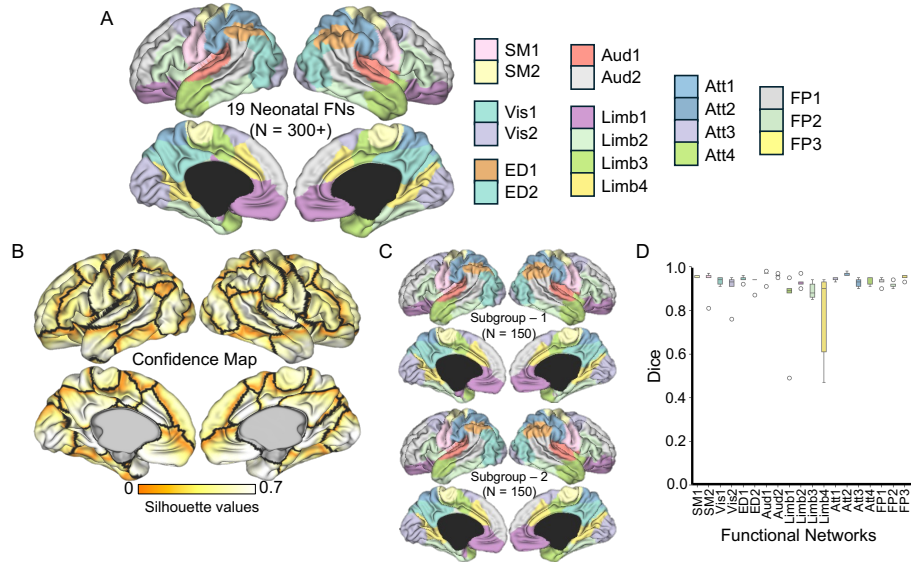


Fig. 4. Fine-grained parcellation of 19 functional networks in neonates derived from a 300-subject dHCP cohort. **A.** Neonatal cortical parcellation of 19 FNs. **B.** Confidence map representing the reliability of the 19 neonatal FNs parcellation. **C.** Discovery and replication of the 19 neonatal FNs cortical parcellation. **D.** Dice score evaluation of FN-wise via bootstrap analysis, assessing the consistency of the parcellation.

2.4 Validation on simulated dataset

Before applying TReND to real-world datasets, we first validated its ability to parcellate FNs using a controlled simulated dataset. The robustness of our

transformer-based autoencoder for feature extraction was tested by keeping the RNMF-KMeans clustering consistent against other conventional mathematically constrained methods, PCA, UMAP and TD. Our approach achieved a Dice of 0.92 and IoU of 0.89, with only 5.82% mislabeled vertices, indicating a $\sim 15\%$ improvement (Fig 2A). To further assess our clustering strategy, we independently validated the RNMF-KMeans approach using features from the TReND framework. The results also demonstrate a $\sim 15\%$ accuracy gain over standalone NMF, kPCA and ICA (Fig 2B). This validation highlights the robustness of our framework in preserving structural patterns while mitigating noise and spurious signals, a key challenge in neonatal fMRI data.

3 Results

3.1 Coarse parcellation of 7 and 19 functional networks in neonates

Building on the validation with the simulated dataset, we applied the TReND framework to the dHCP cohort. Our cluster stability analyses identify 7 and 19 FNs as optimal starting points for delineation of neonatal cortex (Fig 3A). The 7 FNs atlas provide a coarse parcellation of developing networks, derived from 300 term neonates in the dHCP cohort (Fig 3B). Confidence maps reveal lower confidence at boundary regions, suggesting further subdivision (Fig 3C). Bootstrap analysis with 100 combinations confirms high reproducibility, with the 7 FNs atlas achieving an average Dice of ~ 0.92 (Fig 3D). Network-wise variability remains low (Standard deviation (SD) ~ 0.05 Dice), except for the attention network, which exhibits slightly higher variation (Fig 3E). We identify emerging default network in neonates, where the absence of frontal and temporal components reflects the incomplete maturation of higher-order FNs compared to adults (Fig 3F) [29]. In contrast, primary networks, such as somatomotor and visual FNs, exhibit great consistency with the adult atlas (Fig 3F). This finding highlights why adult parcellation is not transferable to neonates, as several higher-order FNs are yet to fully develop.

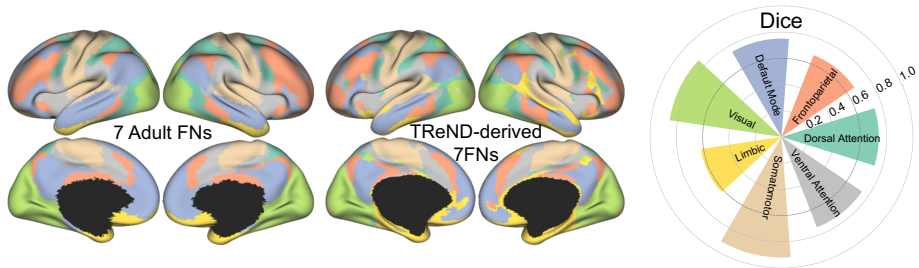


Fig. 5. Validation of TReND’s performance on the HCP-YA dataset. Consistency of the 7 FNs identified by TReND compared to the Yeo 7-network atlas evaluated with dice.

Similarly, the 19 FNs atlas provide a finer-grained parcellation (Fig 4A), further fractionating low-confidence boundary regions from the 7 FNs, reflecting the hierarchical organization of neonatal FC (Fig 4B). This atlas also achieves an average Dice of ~ 0.90 , maintaining strong reproducibility despite increased granularity (Fig 4C). Most networks exhibit an SD of ~ 0.09 Dice, except for the limbic 4 network, which shows higher variability (Fig 4D).

3.2 Cross-dataset performance of TReND using HCP-YA

Although the TReND framework is developed for the neonatal cortex parcellation, we evaluated its applicability on the HCP-YA dataset using the Yeo 7-network atlas as ground truth [3]. All networks, except the limbic network, achieved a Dice score of ~ 0.7 or higher, with the somatomotor and visual networks reaching ~ 0.9 (Fig 5). These variations may have arisen due to TReND’s voxel-wise rather than ROI-wise approach and cohort differences.

4 Conclusion

Our study introduced TReND, a novel transformer-autoencoder framework with RNMF for neonatal FN parcellation. Given that adult FNs are not directly transferable to neonates due to their incomplete maturation, our approach provides a dedicated solution for neonatal brain functional organization. The resulting 7- and 19-network parcellations aim to establish standardized neonatal functional atlases, paving the way for future research in early brain development.

Acknowledgments. This study is funded by NIH R01MH092535, R01MH125333, R01EB031284, R01MH129981, R21MH123930 and P50HD105354.

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

References

1. Behrens, T.E., Johansen-Berg, H., Woolrich, M.W., Smith, S.M., Wheeler King-shott, C.A., Boulby, P.A., Barker, G.J., Sillery, E.L., Sheehan, K., Ciccarelli, O., Thompson, A.J., Brady, J.M., Matthews, P.M.: Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nature Neuroscience* **6**(7), 750–757 (2003)
2. Huang, H., Zhang, J., Jiang, H., Wakana, S., Poetscher, L., Miller, M.I., van Zijl, P.C., Hillis, A.E., Wytik, R., Mori, S.: DTI tractography based parcellation of white matter: application to the mid-sagittal morphology of corpus callosum. *Neuroimage* **26**(1), 195–205 (2005).
3. Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zollei, L., Polimeni, J.R., Fischl, B., Liu, H., Buckner, R.L.: The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*. **106**(3), 1125–1165 (2011).

4. Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., Smith, S.M., Van Essen, D.C.: A multi-modal parcellation of human cerebral cortex. *Nature*. **536**(7615), 171–178 (2016).
5. Peng, Q., Ouyang, M., Wang, J., Yu, Q., Zhao, C., Slinger, M., Li, H., Fan, Y., Hong, B., Huang, H.: Regularized-Ncut: Robust and homogeneous functional parcellation of neonate and adult brain networks. *Artificial Intelligence in Medicine*. **106**, 101872 (2020).
6. Wang, F., Zhang, H., Wu, Z., Hu, D., Zhou, Z., Girault, J.B., Wang, L., Lin, W. and Li, G.: Fine-grained functional parcellation maps of the infant cerebral cortex. *eLife*. **12**, e75401 (2023).
7. Wen, X., Zhao, Y., Chen, G., Zhang, H. and Zhang, D.: Constructing fine-grained spatiotemporal neonatal functional atlases with spectral functional network learning. *Human Brain Mapping*, **45**(8), e26718. (2024)
8. Cao, M., He, Y., Dai, Z., Liao, X., Jeon, T., Ouyang, M., Chalak, L., Bi, Y., Rollins, N., Dong, Q.: Early development of functional network segregation revealed by connectomic analysis of the preterm human brain. *Cerebral Cortex*. **27**(3), 1949–1963 (2017).
9. Xia, Y., Zhao, J., Xu, Y., Duan, D., Xia, M., Jeon, T., Ouyang, M., Chalak, L., Rollins, N., Huang, H.: Development of sensorimotor-visual connectome gradient at birth predicts neurocognitive outcomes at 2 years of age. *iScience*. **27**(2) (2024).
10. Xu, Y., Cao, M., Liao, X., Xia, M., Wang, X., Jeon, T., Ouyang, M., Chalak, L., Rollins, N., Huang, H.: Development and emergence of individual variability in the functional connectivity architecture of the preterm human brain. *Cerebral Cortex*. **29**(10), 4208–4222 (2019).
11. Bassett, D.S., Bullmore, E., Verchinski, B.A., Mattay, V.S., Weinberger, D.R., Meyer-Lindenberg, A.: Hierarchical organization of human cortical networks in health and schizophrenia. *Journal of Neuroscience*. **28**(37), 9239–9248 (2008).
12. Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F.: Correspondence of the brain’s functional architecture during activation and rest. *Proceedings of the National Academy of Sciences*. **106**(31), 13040–13045 (2009).
13. Yerys, B.E., Herrington, J.D., Satterthwaite, T.D., Guy, L., Schultz, R.T., Bassett, D.S.: Globally weaker and topologically different: resting-state connectivity in youth with autism. *Molecular Autism*. **8**, 39 (2017).
14. Oishi K., Chang L., Huang H.: Baby brain atlases. *Neuroimage*. **185**, 865–880 (2019).
15. Huang, H., Huppi, P.S., He, Y., Lerch, J.: Imaging baby brain development. *Neuroimage*. **203**, 116142 (2019).
16. Xu, J., Yang, Y., Huang, D., Gururajapathy, S.S., Ke, Y., Qiao, M., Wang, A., Kumar, H., McGeown, J. and Kwon, E.: Data-driven network neuroscience: On data collection and benchmark. *Advances in Neural Information Processing Systems*, **36**, 21841–21856 (2023).
17. Abrol, A., Fu, Z., Salman, M., Silva, R., Du, Y., Plis, S. and Calhoun, V.: Deep learning encodes robust discriminative neuroimaging representations to outperform standard machine learning. *Nature Communications*, **12**(1), 353 (2021).
18. Tu, J.C., Kim, J.H., Luckett, P., Adeyemo, B., Shimony, J.S., Elison, J.T., Eggebrecht, A.T. and Wheelock, M.D.: Deep-learning based Embedding of Functional Connectivity Profiles for Precision Functional Mapping. *bioRxiv* (2025).

19. Erhardt, E.B., Allen, E.A., Wei, Y., Eichele, T., Calhoun, V.D.: SimTB, a simulation toolbox for fMRI data under a model of spatiotemporal separability. *Neuroimage*. **59**(4), 4160–4167 (2012).
20. Edwards, A.D., Rueckert, D., Smith, S.M., et al.: The Developing Human Connectome Project Neonatal Data Release. *Frontiers in Neuroscience*. **16**, 886772 (2022).
21. Fitzgibbon, S.P., Harrison, S.J., Jenkinson, M., et al.: The Developing Human Connectome Project (dHCP) automated resting-state functional processing framework for newborn infants. *Neuroimage*. **223**, 117303 (2020).
22. Van Essen, D.C., Smith, S.M., Barch, D.M., et al.: The WU-Minn human connectome project: an overview. *Neuroimage*. **80**, 62–79 (2013).
23. Glasser, M.F., Sotiropoulos, S.N., et al.: The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage*. **80**, 105–124 (2013).
24. Vaswani, A., Guyon, I., Luxburg, U.V., Bengio, S., et al.: Attention is all you need. *Advances in Neural Information Processing Systems*, 5998–6008. (2017)
25. Paatero, P., Tapper, U.: Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*. **5**(2), 111–126 (1994).
26. Boutsidis, C., Gallopoulos, E.: SVD based initialization: A head start for nonnegative matrix factorization. *Pattern Recognition*. **41**(4), 1350–1362 (2008).
27. Ben-Hur, A., Elisseeff, A., Guyon, I.: A stability-based method for discovering structure in clustered data. *Biocomputing*. **2002**, 6–17 (2001).
28. Rousseeuw, P.J.: Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*. **20**, 53–65 (1987).
29. Yu, Q., Ouyang, M., Detre, J., Kang, H., Hu, D., Hong, B., Fang, F., Peng, Y. and Huang, H.: Infant brain regional cerebral blood flow increases supporting emergence of the default-mode network. *eLife*. **12**, e78397 (2023)