Dynamic Neural Fields for Learning Atlases of 4D Fetal MRI Time-series

Zeen Chi[1,2*] **Zhongxiao Cong**[1,2*] **Clinton J. Wang**[2] **Yingcheng Liu**[2]
Esra Abaci Turk[3,4] **P. Ellen Grant**[3,4] **S. Mazdak Abulnaga**[2,4,5]
Polina Golland[2] **Neel Dey**[2]
[1]School of Information Science and Technology, ShanghaiTech University [2]MIT CSAIL
[3]Fetal-Neonatal Neuroimaging & Developmental Science Center, Boston Children's Hospital
[4] Harvard Medical School [5]Massachusetts General Hospital
{chize,congzhx}@shanghaitech.edu.cn {dey,abulnaga,liuyc}@mit.edu
{clinton,polina}@csail.mit.edu
{esra.abaciturk,ellen.grant}@childrens.harvard.edu

Abstract

We present a method for fast biomedical image atlas construction using neural fields. Atlases are key to biomedical image analysis tasks, yet conventional and deep network estimation methods remain time-intensive. In this preliminary work, we frame subject-specific atlas building as learning a neural field of deformable spatiotemporal observations. We apply our method to learning subject-specific atlases and motion stabilization of dynamic BOLD MRI time series of fetuses *in utero*. Our method yields high-quality atlases of fetal BOLD time-series with $\sim$5-7$\times$ faster convergence compared to existing work. While our method slightly underperforms well-tuned baselines in terms of anatomical overlap, it estimates templates significantly faster, thus enabling rapid processing and stabilization of large databases of 4D dynamic MRI acquisitions. Code is available at https://github.com/Kidrauh/neural-atlasing.

1 Introduction

Given biomedical image observations, constructing image atlases enables morphometric analyses and registration to a common coordinate system. Current conventional [6, 14, 16, 24, 26, 27] and deep learning methods [9–11, 31, 32] for atlas building yield high-quality atlases with accurate registration at the cost of significant computation time. These computational costs compound further when given *subject-specific* image time-series (e.g., longitudinal repeats) where a new atlas must be constructed for each subject to enable motion stabilization and standardized analyses.

In the context of fetal image analysis, *in-utero* BOLD MRI time series can track changes in fetal and placental oxygenation under induced maternal hyperoxia to identify dysfunction and monitor fetal and maternal well-being [2, 15, 25, 29]. However, the inter-timepoint motion caused by fetal movement and maternal breathing necessitates nonlinear registration of the time series to a common coordinate system for each individual subject to stabilize motion prior to any analysis. To that end, this work presents a method for fast subject-specific spatiotemporal atlasing.

We formulate atlas estimation as the learning of compactly-parameterized dynamic neural fields [20, 21, 23, 28] to represent both the atlas and image-to-atlas deformations. Using our proposed neural representation and training strategy, we rapidly construct high-fidelity subject-specific atlases and stabilize the motion present in BOLD MR images of fetuses *in utero* to enable improved analyses of key BOLD time series-based fetal and maternal biomarkers [29].

*Equal contribution. Work done while visiting MIT CSAIL.

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

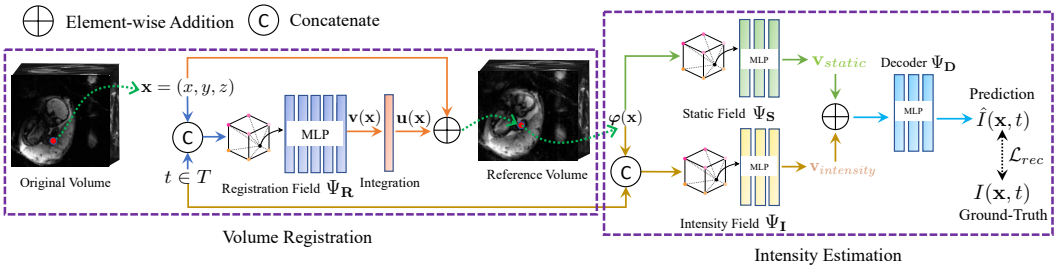

Figure 1: **Architecture.** Our method constructs neural fields for volume registration and intensity estimation, which warp observations to an atlas space and learn the atlas parameters, respectively.

Table 1: **Quantitative results** of baseline comparisons (**top**) and ablations (**bottom**) studying registration performance (via local normalized cross-correlation and weighted dice), deformation qualities (via deformation magnitude, avg. Jacobian determinant, and folding ratio), and runtimes.

	LNCC (↑)	**Wt. Dice** (↑)	$\|\mathbf{u}(\mathbf{x})\|_2$ (↓)	$\lvert J_{\boldsymbol{\varphi}}\rvert$	**% folds** (↓)	**Runtime** (↓)
Unaligned	0.392(0.073)	0.80(0.05)	-	-	-	-
SyGN [7]	0.528(0.075)	**0.91**(0.02)	0.0227(0.0035)	1.000(0.000)	**0**	12hrs / 96-core CPU
AtlasMorph [10]	0.531(0.079)	0.90(0.02)	**0.0083**(0.0014)	1.004(0.003)	**0**	16hrs / A6000 GPU
Ours	**0.579**(0.081)	0.88(0.02)	0.0183(0.0067)	1.004(0.013)	0.01(0.01)	**2.2**hrs / A6000 GPU
(- SVF)	0.503(0.081)	0.85(0.04)	0.0096(0.0021)	1.006(0.010)	0.04(0.02)	1.1hrs / A6000 GPU
(- Divergence)	0.579(0.078)	0.87(0.02)	0.0200(0.0063)	1.013(0.012)	0.06(0.04)	1.5hrs / A6000 GPU
(- Intensity field)	0.578(0.083)	0.88(0.02)	0.0209(0.0086)	1.000(0.018)	0.01(0.01)	2.2hrs / A6000 GPU

2 Methods

Learning Neural Fields. Fig. 1 presents our method consisting of networks for image-to-atlas deformation and atlas estimation. We use three neural fields, each parameterized as a multi-resolution hash encoding followed by a small MLP [19] for efficient processing. We further use stationary velocity fields (SVF) to ensure diffeomorphic deformations [3, 4, 17]. The atlas is produced by an encoder-decoder where the encoder consists of time-invariant (*static*) and time-variant (*intensity*) functions that allow small changes in atlas appearance to account for subtle topological changes.

Given spatial $\mathbf{x} = (x, y, z)$ and temporal $t \in T$ coordinates, the registration field $\Psi_\mathbf{R} : \mathbb{R}^4 \mapsto \mathbb{R}^3$ computes velocities $\mathbf{v}(\mathbf{x})$ which integrate to yield a diffeomorphic displacement field $\mathbf{u}(\mathbf{x})$ between an image at time t and the atlas, such that the deformation between them is $\boldsymbol{\varphi}(\mathbf{x}) = \mathbf{u}(\mathbf{x}) + \mathbf{x}$. On warping the image coordinates into the atlas space, we query $\boldsymbol{\varphi}(\mathbf{x})$ from the static field $\Psi_\mathbf{S} : \mathbb{R}^3 \mapsto \mathbb{R}^n$ to get the feature vector $\mathbf{v}_{static} \in \mathbb{R}^n$ encoding time-invariant latent atlas features. We then query $(\boldsymbol{\varphi}(\mathbf{x}), t)$ from an intensity field $\Psi_\mathbf{I} : \mathbb{R}^4 \mapsto \mathbb{R}^n$ that yields $\mathbf{v}_{intensity} \in \mathbb{R}^n$ encoding the latent intensity differences between $\boldsymbol{\varphi}(\mathbf{x})$ in the atlas and $\mathbf{x}$ in the original image. An MLP $\Psi_\mathbf{D} : \mathbb{R}^n \mapsto \mathbb{R}$ then decodes the fused latent features and yields the estimated intensity $\hat{I}(\mathbf{x}, t)$ of the original image.

Losses. We use the L_1 reconstruction objective $\mathcal{L}_{rec} = \frac{1}{|\Omega|} \sum_{\mathbf{x} \in \Omega} |I(\mathbf{x}, t) - \hat{I}(\mathbf{x}, t)|$ where Ω is the spatial coordinates and I and $\hat{I}$ are ground truth and estimated intensities of the image, respectively. To encourage smooth, locally-rigid, and central deformations, we develop the regularizer $\mathcal{L}_{def} = \lambda_1 \frac{1}{|\Omega|} \sum_{\mathbf{x} \in \Omega} \|\mathbf{u}(\mathbf{x})\|_2 + \lambda_2 \mathcal{L}_{div} + \lambda_3 \|\bar{\mathbf{u}}(\mathbf{x})\|_2^2$, where $\bar{\mathbf{u}}(\mathbf{x})$ is the moving average of displacement vectors [10] and $\mathcal{L}_{div} = \frac{1}{|\Omega|} \sum_{\mathbf{x} \in \Omega} |\mathrm{div}(\mathbf{u}(\mathbf{x}))|^2$ is the divergence loss [30] that encourages locally-rigid deformations which are essential to properly model fetal motion. To reduce folds in the computed deformations, we use the negative Jacobian loss $\mathcal{L}_{jac}$ [18], which reduces the number of negative elements in the determinant of the Jacobian of the deformation. For intensity estimation, we use L_1 regularization $\mathcal{L}_{int}$ on $\mathbf{v}_{intensity}$ to limit temporal appearance changes, and use total variation regularization $\mathcal{L}_{tv} = \mathrm{tv}(\mathbf{v}_{static}) + \mathrm{tv}(\mathbf{v}_{intensity})$ on $\mathbf{v}_{static}$ and $\mathbf{v}_{intensity}$ to encourage piecewise-constant and sharp-edged atlases both spatially and temporally. Our overall objective is $\mathcal{L}(F) = \mathcal{L}_{rec} + \mathcal{L}_{def} + \lambda_{jac}\mathcal{L}_{jac} + \lambda_{int}\mathcal{L}_{int} + \lambda_{tv}\mathcal{L}_{tv}$ where $\lambda_1 = 10^{-3}, \lambda_2 = 5 \times 10^{-4}, \lambda_3 = 0.1, \lambda_{jac} = 1, \lambda_{int} = 0.05$, and $\lambda_{tv} = 0.1$, chosen via grid search on two validation subjects.

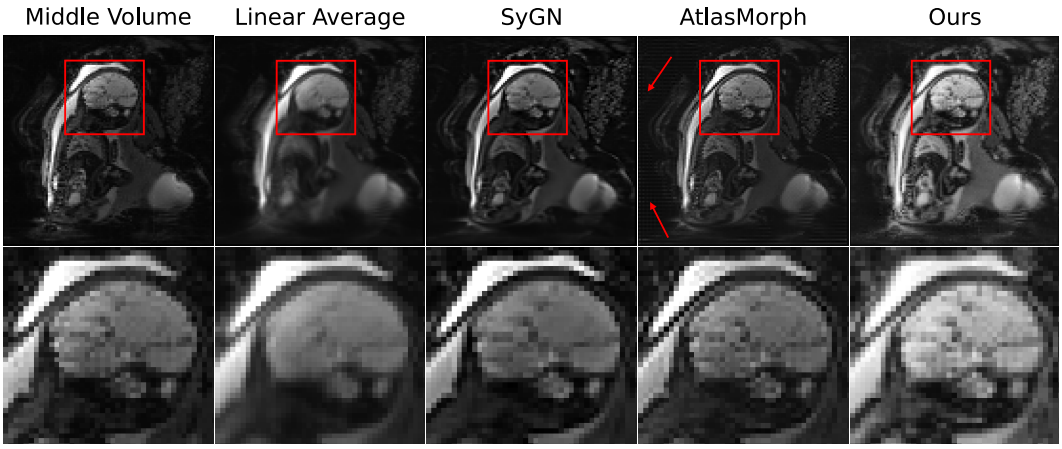

Figure 2: Given an arbitrarily chosen subject, we illustrate the mid-timepoint of the time-series, the temporal linear average, and fetal atlases produced by SyGN [7], AtlasMorph [10], and our method. Atlasmorph creates undesirable checkerboard artifacts (indicated by red arrows).

Atlas Inference. To construct the final atlas (the single time-invariant template) representing the entire time-series, we directly query $(\mathbf{x}, t)$ from the trained atlas encoder-decoder network (Fig. 1 right, intensity estimation). We first calculate the static feature vector $\mathbf{v}_{static}$ and the intensity feature vectors $\mathbf{v}_{intensity}$ at each time step t and then decode $\mathbf{v}_{static} + \frac{1}{T}\sum_{t=1}^{T} \mathbf{v}_{intensity}$ using $\Psi_{\mathbf{D}}$.

3 Experiments

Data and Baselines. We use 11 dynamic BOLD MRI time-series of *in utero* fetal subjects (2 for tuning hyperparameters and modeling decisions and 9 for held-out testing) with a time-series length of 78 to 146 time points per subject [1]. Due to fetal motion and maternal breathing, there is a need for the registration of all images to a common unbiased subject-specific representation. Each image is resampled to $112 \times 112 \times 80$ at $3mm^3$ isotropic resolution. As we use an intensity-based reconstruction loss, we use adaptive histogram equalization [22] for inputs to our model to balance contributions from bright and dark BOLD regions such as the amniotic fluid and fetal body, respectively. We use SyGN [7] and AtlasMorph [10] as representative conventional and deep network baselines, with local normalized cross-correlation (LNCC) [5] as a registration loss which is locally-adaptive and intensity scale-invariant by design. AtlasMorph and our method are trained on a single RTX A6000 GPU and SyGN is optimized on a server CPU using 96 hyperthreaded cores.

Evaluation. Atlas building evaluation is subtle and involves trade-offs between registration accuracy, deformation quality, and runtime [11]. To measure performance, we follow [13] and randomly select 50 MRI pairs for each subject and compose image-to-atlas and atlas-to-image warps to calculate LNCC and multi-region Dice coefficients [12]. Our segmentation labels correspond to the placenta, amniotic fluid, fetal body, fetal brain, and fetal eyes and are generated by an in-house segmentation network. To assess deformation quality, we calculate the average displacement L_2 norm between the atlas and images with a lower value indicating improved template centrality, the mean determinant of the Jacobian matrix $J_{\boldsymbol{\varphi}}(p)$ w.r.t. the input voxel p, and the ratio of deformation folds.

Results. Table 1 reports LNCC and the weighted average Dice scores and deformation statistics comparisons between the baselines and our model. All methods produce invertible deformations. The proposed model achieves best-in-class LNCC but lags behind slightly in terms of Dice score (i.e., anatomical overlap). In terms of runtime, our proposed model converges $5.5 - 7.4\times$ faster than baselines yielding high-fidelity templates (see Fig. 2) with smooth and invertible deformations. However, if the tuned baselines are optimized to convergence, they currently yield improved anatomical overlap. Ablations removing the SVF formulation, the divergence loss, and $\Psi_{\mathbf{I}}$ all worsen performance.

4 Conclusions and Future Directions

We demonstrate that dynamic neural fields learn atlases of 4D fetal BOLD MRI time-series significantly faster than current methods. These speed gains are especially relevant to subject-specific atlas building of large collections of subjects imaged using 4D dynamic MRI. Currently, our preliminary work finds that well-tuned baselines optimized for longer still achieve better registration overlap in terms of Dice. This performance gap points to several future directions: (1) Fetal BOLD MRI time series are temporally sampled at only $\sim$0.28 frames per second (FPS) as compared to conventional video (24+ FPS) for which existing work on dynamic neural fields was developed. This gives rise to large, erratic motion between consecutive timepoints, and may require modification to existing positional encoding functions which assume temporal smoothness. (2) High-performing mono-modal biomedical image registration frameworks typically use LNCC [8] as a registration loss. However, due to the scale and shift-invariant formulation of LNCC, neural regression networks trained with LNCC require significant regularization to guide them towards non-degenerate solutions, which we find can introduce significant artifacts in the estimated atlas. Future work may seek to mitigate this tradeoff by constraining the optimization space of the network or using data-driven priors.

5 Acknowledgements

We gratefully acknowledge funding from NIH NIBIB 5R01EB032708, NIH NICHD R01HD100009, NIH NIA 5R01AG064027, and NIH NIA 5R01AG070988. We thank all the participants for providing the data in the study.

References

[1] S. M. Abulnaga, S. I. Young, K. Hobgood, E. Pan, C. J. Wang, P. E. Grant, E. Abaci Turk, and P. Golland. Automatic segmentation of the placenta in bold mri time series. In *Perinatal, Preterm and Paediatric Image Analysis: 7th International Workshop, PIPPI 2022, Held in Conjunction with MICCAI 2022, Singapore, September 18, 2022, Proceedings*, pages 1–12. Springer, 2022.

[2] S. Aimot-Macron, L. Salomon, B. Deloison, R. Thiam, C. Cuenod, O. Clement, and N. Siauve. In vivo mri assessment of placental and foetal oxygenation changes in a rat model of growth restriction using blood oxygen level-dependent (bold) magnetic resonance imaging. *European radiology*, 23:1335–1342, 2013.

[3] V. Arsigny, O. Commowick, X. Pennec, and N. Ayache. A log-euclidean framework for statistics on diffeomorphisms. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006: 9th International Conference, Copenhagen, Denmark, October 1-6, 2006. Proceedings, Part I 9*, pages 924–931. Springer, 2006.

[4] J. Ashburner. A fast diffeomorphic image registration algorithm. *Neuroimage*, 38(1):95–113, 2007.

[5] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis*, 12(1):26–41, 2008.

[6] B. B. Avants, N. Tustison, G. Song, et al. Advanced normalization tools (ants). *Insight j*, 2(365):1–35, 2009.

[7] B. B. Avants, P. Yushkevich, J. Pluta, D. Minkoff, M. Korczykowski, J. Detre, and J. C. Gee. The optimal template effect in hippocampus studies of diseased populations. *Neuroimage*, 49(3):2457–2466, 2010.

[8] B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, and J. C. Gee. A reproducible evaluation of ants similarity metric performance in brain image registration. *Neuroimage*, 54(3):2033–2044, 2011.

[9] L. Chen, Z. Wu, D. Hu, Y. Pei, F. Zhao, Y. Sun, Y. Wang, W. Lin, L. Wang, G. Li, et al. Construction of longitudinally consistent 4d infant cerebellum atlases based on deep learning. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part IV 24*, pages 139–149. Springer, 2021.

[10] A. V. Dalca, M. Rakic, J. Guttag, and M. R. Sabuncu. Learning conditional deformable templates with convolutional networks. *NeurIPS: Neural Information Processing Systems*, 2019.

[11] N. Dey, M. Ren, A. V. Dalca, and G. Gerig. Generative adversarial registration for improved conditional deformable templates. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 3929–3941, 2021.

[12] L. R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.

[13] Z. Ding and M. Niethammer. Aladdin: Joint atlas building and diffeomorphic registration learning with pairwise alignment. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 20784–20793, 2022.

[14] M. Kuklisova-Murgasova, P. Aljabar, L. Srinivasan, S. J. Counsell, V. Doria, A. Serag, I. S. Gousias, J. P. Boardman, M. A. Rutherford, A. D. Edwards, et al. A dynamic 4d probabilistic atlas of the developing brain. *NeuroImage*, 54(4):2750–2763, 2011.

[15] J. Luo, E. A. Turk, T. Hahn, M. Teulon Gonzalez, B. Gagoski, C. Bibbo, A. Palanisamy, C. Tempany, A. Torrado-Carvajal, N. Malpica, et al. Human placental and fetal response to maternal hyperoxygenation in iugr pregnancy as measured by bold mri. In *Proceedings of the 23rd Annual Meeting of ISMRM, Toronto, Ontario, Canada*, page 633, 2015.

[16] A. Makropoulos, P. Aljabar, R. Wright, B. Hüning, N. Merchant, T. Arichi, N. Tusor, J. V. Hajnal, A. D. Edwards, S. J. Counsell, et al. Regional growth and atlasing of the developing human brain. *Neuroimage*, 125:456–478, 2016.

[17] M. Modat, P. Daga, M. J. Cardoso, S. Ourselin, G. R. Ridgway, and J. Ashburner. Parametric non-rigid registration using a stationary velocity field. In *2012 IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 145–150. IEEE, 2012.

[18] T. C. Mok and A. C. Chung. Large deformation diffeomorphic image registration with laplacian pyramid networks. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part III 23*, pages 211–221. Springer, 2020.

[19] T. Müller, A. Evans, C. Schied, and A. Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics (ToG)*, 41(4):1–15, 2022.

[20] K. Park, U. Sinha, J. T. Barron, S. Bouaziz, D. B. Goldman, S. M. Seitz, and R. Martin-Brualla. Nerfies: Deformable neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5865–5874, 2021.

[21] K. Park, U. Sinha, P. Hedman, J. T. Barron, S. Bouaziz, D. B. Goldman, R. Martin-Brualla, and S. M. Seitz. Hypernerf: A higher-dimensional representation for topologically varying neural radiance fields. *arXiv preprint arXiv:2106.13228*, 2021.

[22] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld. Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*, 39(3):355–368, 1987.

[23] A. Pumarola, E. Corona, G. Pons-Moll, and F. Moreno-Noguer. D-nerf: Neural radiance fields for dynamic scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10318–10327, 2021.

[24] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes. Nonrigid registration using free-form deformations: application to breast mr images. *IEEE transactions on medical imaging*, 18 (8):712–721, 1999.

[25] V. Schöpf, G. Kasprian, P. C. Brugger, and D. Prayer. Watching the fetal brain at 'rest'. *International Journal of Developmental Neuroscience*, 30(1):11–17, 2012.

[26] A. Schuh, M. Murgasova, A. Makropoulos, C. Ledig, S. J. Counsell, J. V. Hajnal, P. Aljabar, and D. Rueckert. Construction of a 4d brain atlas and growth model using diffeomorphic registration. In *Spatio-temporal Image Analysis for Longitudinal and Time-Series Image Data: Third International Workshop, STIA 2014, Held in Conjunction with MICCAI 2014, Boston, MA, USA, September 18, 2014, Revised Selected Papers 3*, pages 27–37. Springer, 2015.

[27] A. Serag, V. Kyriakopoulou, M. A. Rutherford, A. D. Edwards, J. V. Hajnal, P. Aljabar, S. J. Counsell, J. Boardman, and D. Rueckert. A multi-channel 4d probabilistic atlas of the developing brain: application to fetuses and neonates. *Annals of the BMVA*, 2012(3):1–14, 2012.

[28] L. Song, A. Chen, Z. Li, Z. Chen, L. Chen, J. Yuan, Y. Xu, and A. Geiger. Nerfplayer: A streamable dynamic scene representation with decomposed neural radiance fields. *arXiv preprint arXiv:2210.15947*, 2022.

[29] A. Sørensen, D. Peters, C. Simonsen, M. Pedersen, B. Stausbøl-Grøn, O. B. Christiansen, G. Lingman, and N. Uldbjerg. Changes in human fetal oxygenation during maternal hyperoxia as estimated by bold mri. *Prenatal diagnosis*, 33(2):141–145, 2013.

[30] E. Tretschk, A. Tewari, V. Golyanik, M. Zollhöfer, C. Lassner, and C. Theobalt. Non-rigid neural radiance fields: Reconstruction and novel view synthesis of a dynamic scene from monocular video. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 12959–12970, 2021.

[31] Y. Zhang, F. Shi, G. Wu, L. Wang, P.-T. Yap, and D. Shen. Consistent spatial-temporal longitudinal atlas construction for developing infant brains. *IEEE transactions on medical imaging*, 35(12):2568–2577, 2016.

[32] F. Zhao, Z. Wu, L. Wang, W. Lin, S. Xia, G. Li, and U. B. C. P. Consortium. Learning 4d infant cortical surface atlas with unsupervised spherical networks. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part II 24*, pages 262–272. Springer, 2021.