Temporal Atlas-Guided Generation of Longitudinal Data via Geometric Latent Embeddings

Shaoju Wu, Jian Wang, Sila Kurugol, and Andy Tsai

Boston Children's Hospital and Harvard Medical School, Boston, MA, USA
shaoju.wu@childrens.harvard.edu

Abstract. The spatiotemporal changes of a developing anatomical structure is a dynamic process, and quantifying this process within a population and between populations is a fundamental yet challenging task in medical image analysis. Central to this task is the availability of longitudinal imaging data for 4D statistical shape analysis. Unfortunately, this type of longitudinal data is expensive, time-consuming, and difficult to collect. Practically, the majority of imaging data are 3D cross-sectional data, which are inadequate in describing the dynamic shape changes of anatomical structures. In this paper, we introduce a novel temporal atlas-guided deep learning model for longitudinal data generation. Unlike existing methods that directly generate longitudinal data from input images or sequences, we characterize distinctive geometric shape representations in both cross-sectional and longitudinal latent spaces of diffeomorphisms, while optimizing the quality of both atlas and longitudinal data generation. To the best of our knowledge, this is the first deep learning approach that leverages temporal atlas-based representation for longitudinal data generation. The innovative nature of our framework lies in its ability to jointly perform within-age and cross-age shape registration, thus maximizing registration performance while maintaining desirable deformation qualities. Our work's ability to model spatiotemporal dynamics makes it highly versatile and applicable to a wide range of domains, including modeling the normal and abnormal development of anatomical structures for improved clinical diagnosis and treatment planning. The code of this work is available at https://github.com/wushaoju/TAG-GLE.

Keywords: Longitudinal Data Generation · Atlas building · Deep Learning

1 Introduction

The spatiotemporal changes of a developing anatomical structure is a dynamic process, influenced by a complex interplay of various factors including genetics, nutrition, environment, and disease [23, 4, 6]. Quantifying this dynamic process

[1] S. Wu and J. Wang – Equal contribution.

within a population and between populations is a fundamental yet challenging task in medical image analysis. Central to this task is longitudinal imaging data, which consist of repeat imaging of the same subject taken at different time points. This type of data is perfectly suited for capturing the spatiotemporal changes of a developing anatomical structure. Investigators have analyzed longitudinal data for a variety of important clinical applications, including tracking the maturation of a child's anatomy [17, 5], determining the progression of disease [15, 14], monitoring organ response to therapy [18, 12], isolating structural differences in congenital anomalies over time [8, 20], and predicting patient outcomes [24, 11]. In recent years, learning-based methods have emerged as powerful tools for modeling 4D spatiotemporal data in clinical contexts. For example, Yoon et al. [29] introduced a sequence-aware diffusion model to generate temporally consistent 3D brain MRIs, while Puglisi et al. [19] used a latent diffusion model trained on longitudinal MRI data to simulate Alzheimer's-related structural changes over time. Despite promising results, these models typically require large-scale longitudinal training datasets, which are rarely available in practice. In contrast, most available medical imaging data are cross-sectional, generated at a single time point from multiple individuals. Cross-sectional imaging data are inadequate in describing the dynamic spatiotemporal changes of developing anatomical structures. They also do not provide information about cause-and-effect relationships, as these data represent only a snapshot in time, encoding information from a single moment. Ideally, to study the complex dynamic nature of a developing structure, we would need subject-specific 3D longitudinal data (i.e., 4D data) to capture the time-varying structural changes. These 4D data encode the inherent correlations and causalities of repeated acquisitions from the same subject, and provide not only population (i.e., average) trajectories of change, but also individual (i.e., subject-specific) profile changes. Furthermore, these 4D data provide improved statistical power over cross-sectional data. In short, longitudinal imaging data is the data type of choice in capturing and understanding the temporal profile of a developing anatomical structure. However, they are expensive, time-consuming, and difficult to collect, hence they are in short supply.

Contribution. We propose a novel deep learning (DL)-based framework that jointly performs two different but related medical image analysis tasks, specifically, atlas building and longitudinal data generation. The goal of atlas building is to map a large number of images onto a common coordinate system, thereby allowing us to study intra-population variabilities and inter-population differences [13, 9, 30]. This task is focused on capturing population trajectories of change. In contrast, the goal of longitudinal data generation is to capture subject-specific trajectories of change. In essence, our model jointly performs within-age shape registration for atlas building, and cross-age shape registration for longitudinal data generation, with the goal of improved spatiotemporal characterization of developing anatomical structures. We believe that by integrating these two tasks within a unified framework, they leverage each other's strengths in improving their individual performances. Fundamental to our proposed framework is the

strategic use of diffeomorphic non-rigid registration to optimize a smooth and invertible transformation between two anatomical shapes [22, 26].

2 Background: Atlas Building

Before introducing our framework, we briefly review the concepts of atlas building in the context of large deformation diffeomorphic metric mapping (LDDMM) [3], and its relationship to deformation-based shape representations. Given a set of images $I_1, \cdots, I_N$ with N being the number of images in the set, the problem of atlas building is to find a mean or template image I and transformations $\phi_1, \cdots, \phi_N$ that minimize the energy function

$$E(I, v_0^n) = \sum_{n=1}^{N} \frac{1}{\sigma^2} \text{Dist}[I \circ \phi_1^{-1}(v_0^n), I_n] + \text{Reg}(v_0^n), \tag{1}$$

where σ^2 is the noise variance in images, and $\circ$ denotes a composition operator. Solving v_0^n and $\phi_1^{-1}(v_0^n)$ involve the geodesic shooting process which states that the geodesic path $\{\psi_t\}$ can be uniquely determined by integrating a given initial velocity v_0 forward in time by using the Euler-Poincaré differential (EPDiff) equation [1, 16] as follows:

$$\frac{\partial v_t}{\partial t} = -K\left[(Dv_t)^T \cdot m_t + Dm_t \cdot v_t + m_t \cdot \text{div}\, v_t\right], \quad \frac{d\phi_t^{-1}}{dt} = -D\phi_t^{-1} \cdot v_t, \tag{2}$$

where the operator D denotes a Jacobian matrix, and $\cdot$ represents element-wise matrix multiplication. Here, div is the divergence, and K is a smoothing operator with its inverse operator given by L (a symmetric and positive-definite differential operator). Note that $L : V \to V^*$ maps a tangent vector $v_t \in V$ into its dual space as a momentum vector $m_t \in V^*$. This is typically denoted as $m_t = Lv_t$ or $v_t = Km_t$.

The atlas-guided shape representations v_0^n efficiently capture the structural variations over different populations and enhance the performance across various image processing tasks [27, 31]. Expanding on this idea, we develop shape representations by building temporal atlases in the latent space, then evolving them as distinctive geometric features for longitudinal data generation. Our longitudinal data generation model effectively preserves anatomical consistency over time despite having sparse cross-sectional data.

3 Methodology

In this section, we introduce our model that learns latent shape embeddings simultaneously in the cross-sectional and longitudinal space. We show the overall network architecture in Fig. 1. We drop the time index t, i.e., $v_0 \triangleq v$, to simplify the notations in the following sections.

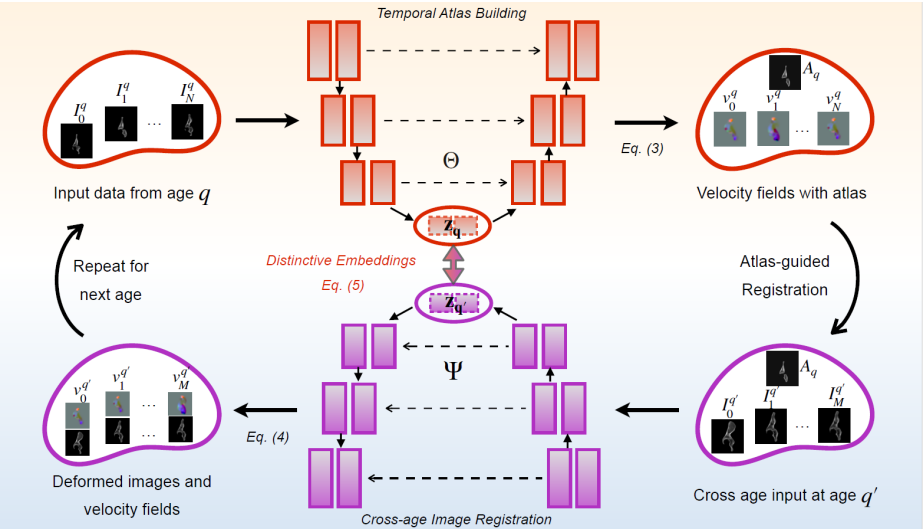

Fig. 1. Overview of our model. **Training**: Our network constructs a sequence of temporal atlases, starting at age q and modeling deformations across target ages. A geometric embedding loss in the latent space enforces structured, biologically meaningful anatomical evolution. Iterative refinement enhances the interplay between cross-sectional and longitudinal dynamics, generating anatomically coherent cross-age data. **Inference**: For each subject, we first align their baseline images to the corresponding age-specific atlas. The trained model Ψ operates as a longitudinal data generator, synthesizing anatomically plausible trajectories for unseen subjects by extracting geometric features through the atlas. It uses generated velocity fields from cross-age registration between age-specific atlases to deform subject-specific structures over time, ensuring a smooth and temporally consistent evolution.

Temporal Atlas Building. Given a set of discrete time points, we aim to construct a temporal atlas that captures the morphological evolution across these time points. Specifically, we build a sequence of atlases, where each atlas characterizes the average anatomical structure at time q. The temporal atlas construction involves optimizing the parameters Θ of a neural network that defines the atlases A_q and their inter-relationships over time. The corresponding temporal atlas loss function is defined as:

$$\mathcal{L}_{\text{tempo-atlas}}(\Theta) = \sum_{q=1}^{Q} \sum_{i=1}^{N_q} \text{Dist}\left[\phi_{q\to i}^{-1}(\Theta) \circ A_q, I_i^q\right] + \text{Reg}(v_i^q), \; s.t. \text{ Eq. (2)}. \quad (3)$$

Cross-age Image Registration. We introduce the network with parameters Ψ to model the longitudinal anatomical changes. Specifically, given that cross-sectional data I_i^q is longitudinally sparse, we align the source atlas A_q to the longitudinal images $\{I_{q'}^m\}_{m=1}^M$ of target age q' using the transformation

$\phi_{q\to q'}$. As such, the longitudinal registration loss is given by:

$$\mathcal{L}_{\text{longitudinal}}(\Psi) = \sum_{m=1}^{M} \text{Dist}[\phi^{-1}_{q\to q'}(\Psi)\circ A_q, I^m_{q'}]^2 + \text{Reg}(v^{q'}_m),\ s.t.\ \text{Eq. (2)}. \quad (4)$$

Distinctive Latent Embeddings. We define the loss between the latent embeddings learned from Eq. (3) and Eq. (4) as:

$$\begin{aligned}\mathcal{L}_{\text{latent}}(\mathbf{Z}) = \mathbb{E}_{(\mathbf{z}^i_q,\mathbf{z}^m_{q'})}\big[-\log(\sigma(\text{sim}(\mathbf{z}^i_q,\mathbf{z}^m_{q'})))\big] \\ + \lambda\,\mathbb{E}_{(\mathbf{z}^i_q,\mathbf{z}^m_{q'})}\left[\text{Corr}\left(\|\mathbf{z}^i_q-\mathbf{z}^m_{q'}\|^2,\ \|v^i_q-v^m_{q'}\|^2\right)\right], \end{aligned} \quad (5)$$

where the first term is the similarity term that computes the expected loss over pairs of embeddings $\mathbf{z}^i_q$ and $\mathbf{z}^m_{q'}$. Here, $\text{sim}(\cdot,\cdot)$ is the cosine similarity, where $\text{sim}(\mathbf{z}_1,\mathbf{z}_2) = \frac{\mathbf{z}_1\cdot\mathbf{z}_2}{\|\mathbf{z}_1\|\|\mathbf{z}_2\|}$, and σ denotes the sigmoid function. Minimizing $-\log(\sigma(.))$ enforces high similarity for related embeddings (e.g., positive pairs), while penalizing low-similarity cases. This encourages related embeddings to cluster tightly in latent space. The second term measures the weighted correlation between the squared Euclidean distance of the embeddings, $\|\mathbf{z}^i_q-\mathbf{z}^m_{q'}\|^2$, and the squared distance of the associated velocity fields, $\|v^i_q-v^m_{q'}\|^2$, thereby ensuring that the latent space preserves the geometric or temporal relationships encoded by the velocity fields. The parameter λ controls the strength of this alignment. This latter term encourages the latent space to capture meaningful geometric or temporal differences indicated by the velocity fields. By employing the distinctive geometric embeddings in the temporal atlas-building and longitudinal registration spaces, we derive our total loss function as:

$$\ell = \mathcal{L}_{\text{tempo-atlas}}(\Theta) + \mathcal{L}_{\text{longitudinal}}(\Psi) + \gamma\,\mathcal{L}_{\text{latent}}(\mathbf{Z}), \quad (6)$$

where γ controls the strength of the latent embeddings alignment. This function enforces within-age shape consistency by clustering anatomically similar structures while preserving between age morphological changes through velocity-guided embedding distances, effectively disentangling progressive shape variations across time.

4 Experiments

To demonstrate the effectiveness of our proposed framework, we conducted both quantitative and qualitative experiments between our model and state-of-the-art (SOTA) techniques in atlas building and longitudinal data generation.

Dataset & Preprocessing. We retrospectively collected cross-sectional and longitudinal computed tomography (CT) datasets of normal male hips for training and evaluation, respectively. For training, we curated a cross-sectional CT dataset from male patients with normal hips acquired between 2011 and 2015. To ensure normality, we excluded patients with (1) hip pain, (2) congenital or developmental musculoskeletal anomalies, and (3) prior hip trauma. This search

yielded 77 CT studies from 77 male patients (mean age: 10.4 ± 3.8 years; range: 4-18 years), totaling 154 hips. For longitudinal analysis, we searched the same archive (2010–2024) and applied the same inclusion and exclusion criteria. We included only patients with at least two CT studies obtained more than one year apart. This search identified 12 male patients with 34 CT studies (mean age: 10.2 ± 3.3 years; range: 4–16 years), totaling 68 hips. We segmented both left and right hips from all CT images using the pre-trained TotalSegmentator [28] and rigidly aligned them to a template hip image space (256^3 voxels, 1 mm resolution). To ensure computational efficiency and maintain consistency across all models, the images were downsampled to 128^3 voxels, allowing for fair and standardized comparisons.

Training Setup. We utilized Geo-SIC [27] as the backbone architecture for training our model in atlas building and longitudinal data generation. The model was trained for 1000 epochs with a batch size of 4, using the Adam optimizer with a learning rate of 0.0005. Training required approximately 4 hours per age on a Quadro RTX A6000 GPU. All implementations were based on PyTorch 2.0.1. Additionally, we performed a sensitivity analysis using a grid search to optimize the registration parameter γ in Eq. (6), and we identified the optimal value as 0.0001.

Atlas Evaluation. To evaluate the quality of the generated temporal hip atlas, we computed the accuracy of atlas-to-patient registration for each specific age (from 4- to 18-years-old) and compared it other SOTA techniques. Specifically, we utilized the mean surface distance (MSD) and the Dice score to compare the results of our atlas building technique to those generated via B-spline [7], Multi-contrast registration (M-c. reg.) [25], VoxelMorph (VM) [2], and our method without the distinctive latent embedding (l.e.) loss.

Longitudinal Data Evaluation. We assessed our longitudinal hip data generator by comparing synthetic hip CTs to real ones using Dice scores. For each patient, a real hip CT was used to generate a synthetic version matching the age of the next real CT in the dataset. The Dice score was computed between the real and synthetic CT scans. For example, given real CTs at ages 8, 10, and 13, we used the 8-year-old CT to generate a synthetic 10-year-old CT, which we then compare to the real 10-year-old CT scan. This process was repeated for subsequent ages. We benchmarked our method against two alternatives: (1) Nearest-Age Hip Approximation (NAHA) with the idea derive from [21], which utilizes the real hip CT that is nearest in age to the target hip CT data to approximate longitudinal changes of the same subject (in essence neglecting age-based changes), and (2) Age-Based Atlas Approximation (ABAA)[10], which relies on general age-specific atlases rather than subject-specific data.

5 Results and Discussion

Fig. 2 showed that our temporal atlas generation method outperformed SOTA atlas building models, producing sharper, more anatomically accurate atlases with fewer artifacts. Other methods often introduced blurry boundaries and

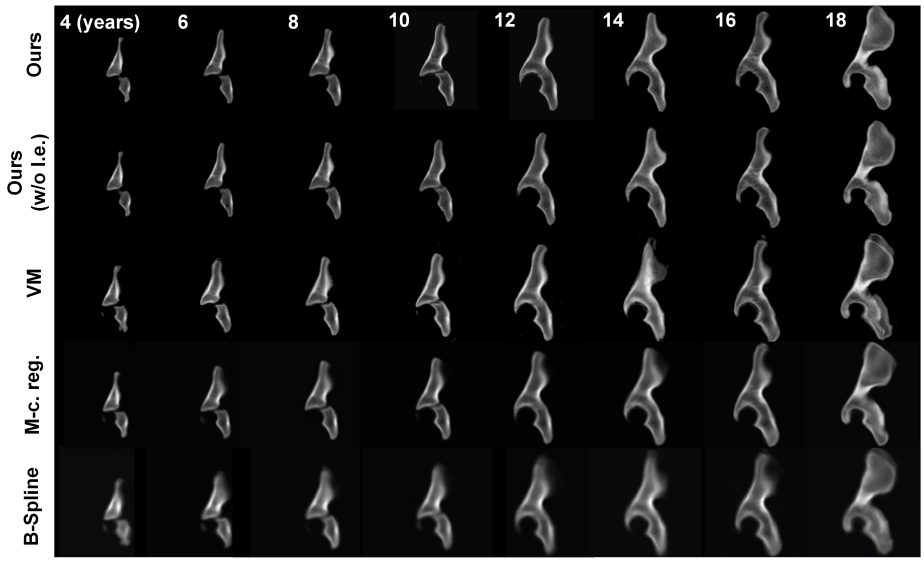

Fig. 2. Comparison of our method with SOTA techniques for temporal hip atlas generation. The age ranged from 4- to 18-years-old. Visually, our method achieved the best performance in atlas building, producing the sharpest boundaries around the growth centers of the hips (i.e., around the tri-radiate cartilages) and the margins of the hip bones (i.e., around the ilium and ischium) compared to other SOTA techniques.

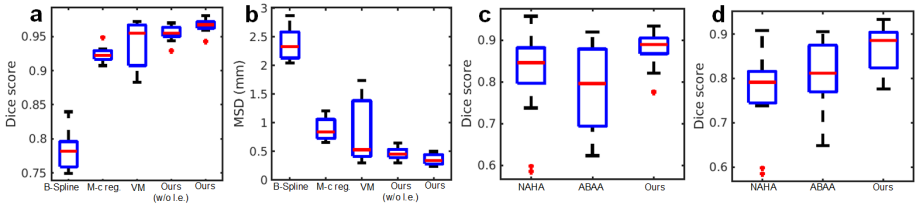

Fig. 3. Boxplots of the atlas-to-patient (within-age) registration results based on the mean Dice score (a) and the MSD (b). Boxplots of the longitudinal data generation results for all testing cases (c) and for the cases when the age gap is >1 year (d) based on mean Dice score. The ages of the subjects in these experiments ranged from 4 to 18 years. Our method achieved superior performance compared to others.

distortions, particularly near the iliac crest, as seen in the 14-year-old temporal atlas where the VM method [2] exhibited substantial artifacts around the posterior iliac crest, while the M-c. reg. [25] and B-Spline [7] methods showed similar but less severe distortions. In contrast, our approach generated atlases with well-defined, artifact-free boundaries, preserving fine anatomical details.

Fig. 3a and Fig. 3b quantitatively demonstrated the superiority of our method, achieving the highest mean Dice score of 0.97 and the lowest mean surface dis-

tance (MSD) of 0.36 mm. These results highlighted our method's robustness in capturing temporal shape variations with high precision, ensuring more accurate spatiotemporal characterization of hip development for clinical and research applications.

Fig. 3c and Fig. 3d showed quantitatively our method's performance for longitudinal data generation across all models. Our proposed method achieved the highest mean Dice score of 0.88, demonstrating superior accuracy in capturing temporal hip shape changes. Notably, the performance variations across baseline methods indicated that age-related factors significantly impact longitudinal shape prediction, particularly for cases with larger age gaps. This further highlighted the robustness of our approach in modeling complex temporal deformations, ensuring more reliable and anatomically consistent predictions over time.

Fig. 4 presented a case study on longitudinal data generation, illustrating the generated images with shape trajectories produced by all models. Our method, guided by the temporal atlas, demonstrated the closest alignment with ground truth hip surfaces, even when predicting shapes with large age gaps (e.g., using a 12-year-old hip CT to predict the corresponding 15-year-old hip CT). This underscored the effectiveness of our method in modeling spatiotemporal hip growth solely from cross-sectional data, generating anatomically consistent and accurate longitudinal data predictions.

6 Conclusion

In this work, we introduced a temporal atlas-guided DL framework to generate longitudinal data using only cross-sectional inputs. Our approach leverages geometric latent spaces of diffeomorphisms to optimize both atlas quality and the fidelity of shape progression. By jointly performing within-age and cross-age registration, our model improves accuracy while preserving anatomically consistent deformations. This framework provides a powerful tool for modeling normal and pathological anatomical development, with applications in clinical diagnosis and disease progression analysis. Additionally, generating high-quality synthetic longitudinal data helps address the scarcity of real datasets, advancing research in medical imaging and computational anatomy. Future work includes integrating patient-specific growth models with biomechanical priors to improve prediction accuracy. Enforcing physics-based constraints and leveraging hybrid statistical-DL models may enhance robustness and personalization.

Acknowledgements. This work was supported in part by National Institute of Child Health and Human Development (No. R21HD108634), National Institute of Diabetic and Digestive and Kidney Diseases (No. R21DK123569 and R01DK125561), and National Institute of Biomedical Imaging and Bioengineering (No. R21EB029627).

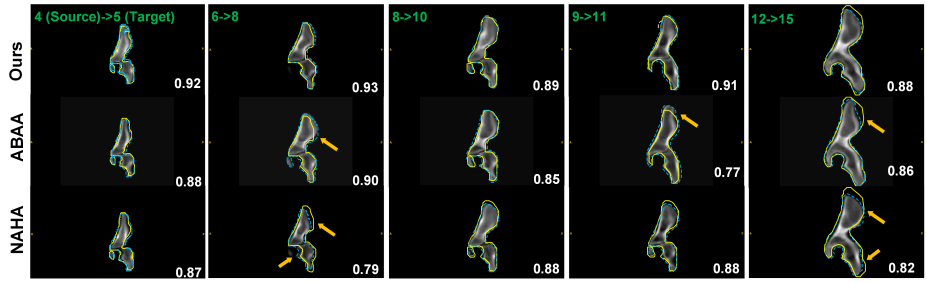

Fig. 4. Comparison of various methods for the longitudinal data generation task based on a real longitudinal dataset. For example, in the first column, given an input hip dataset from a 4-year-old subject (source), the longitudinal data generator predicted the 3D synthetic longitudinal CT hip shape of the same subject at the age of 5. The source and target ages are denoted in green text in the left upper corner of each frame. The segmentation contour (blue dashed line) produced by our method aligns better with the ground truth contour (yellow line) than other techniques, as indicated by the higher Dice scores (white text). Arrows highlight regions with large alignment errors.

Disclosure of Interests. The authors declare that they have no conflict of interest.

References

1. Arnold, V.: Sur la géométrie différentielle des groupes de lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. In: Annales de l'institut Fourier. vol. 16, pp. 319–361 (1966)
2. Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J., Dalca, A.V.: Voxelmorph: a learning framework for deformable medical image registration. IEEE transactions on medical imaging **38**(8), 1788–1800 (2019)
3. Beg, M.F., Miller, M.I., Trouvé, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. International journal of computer vision **61**, 139–157 (2005)
4. Bersvendsen, J., Toews, M., Danudibroto, A., Wells III, W.M., Urheim, S., Estépar, R.S.J., Samset, E.: Robust spatio-temporal registration of 4d cardiac ultrasound sequences. In: Medical Imaging 2016: Ultrasonic Imaging and Tomography. vol. 9790, pp. 122–128. SPIE (2016)
5. Canada, K.L., Hancock, G.R., Riggins, T.: Modeling longitudinal changes in hippocampal subfields and relations with memory from early-to mid-childhood. Developmental Cognitive Neuroscience **48**, 100947 (2021)
6. Chi, Z., Cong, Z., Wang, C.J., Liu, Y., Turk, E.A., Grant, P.E., Abulnaga, S.M., Golland, P., Dey, N.: Dynamic neural fields for learning atlases of 4d fetal mri time-series. arXiv preprint arXiv:2311.02874 (2023)
7. De Vos, B.D., Berendsen, F.F., Viergever, M.A., Sokooti, H., Staring, M., Išgum, I.: A deep learning framework for unsupervised affine and deformable image registration. Medical image analysis **52**, 128–143 (2019)

8. Emam, D., Aertsen, M., Van der Veeken, L., Fidon, L., Patkee, P., Kyriakopoulou, V., De Catte, L., Russo, F., Demaerel, P., Vercauteren, T., et al.: Longitudinal mri evaluation of brain development in fetuses with congenital diaphragmatic hernia around the time of fetal endotracheal occlusion. American Journal of Neuroradiology **44**(2), 205–211 (2023)
9. Fletcher, P.T., Venkatasubramanian, S., Joshi, S.: The geometric median on riemannian manifolds with application to robust atlas estimation. NeuroImage **45**(1), S143–S152 (2009)
10. Fonov, V., Evans, A.C., Botteron, K., Almli, C.R., McKinstry, R.C., Collins, D.L., Group, B.D.C., et al.: Unbiased average age-appropriate atlases for pediatric studies. Neuroimage **54**(1), 313–327 (2011)
11. Holste, G., Lin, M., Zhou, R., Wang, F., Liu, L., Yan, Q., Van Tassel, S.H., Kovacs, K., Chew, E.Y., Lu, Z., et al.: Harnessing the power of longitudinal medical imaging for eye disease prognosis using transformer-based sequence modeling. NPJ Digital Medicine **7**(1), 216 (2024)
12. Jin, C., Yu, H., Ke, J., Ding, P., Yi, Y., Jiang, X., Duan, X., Tang, J., Chang, D.T., Wu, X., et al.: Predicting treatment response from longitudinal images using multi-task deep learning. Nature communications **12**(1), 1851 (2021)
13. Joshi, S., Davis, B., Jomier, M., Gerig, G.: Unbiased diffeomorphic atlas construction for computational anatomy. NeuroImage **23**, S151–S160 (2004)
14. Liu, C., Hu, F., Li, Z., Wang, Y., Zhang, X.: Anterior pelvic plane: a potentially useful pelvic anatomical reference plane in assessing the patients' ideal pelvic parameters without the influence of spinal sagittal deformity. Global Spine Journal **12**(4), 567–572 (2022)
15. Mehta, S., Gajjar, S.R., Padgett, K.R., Asher, D., Stoyanova, R., Ford, J.C., Mellon, E.A.: Daily tracking of glioblastoma resection cavity, cerebral edema, and tumor volume with mri-guided radiation therapy. Cureus **10**(3) (2018)
16. Miller, M.I., Trouvé, A., Younes, L.: Geodesic shooting for computational anatomy. Journal of mathematical imaging and vision **24**, 209–228 (2006)
17. Mills, K.L., Tamnes, C.K.: Methods and considerations for longitudinal structural brain imaging analysis across development. Developmental cognitive neuroscience **9**, 172–190 (2014)
18. Nagtegaal, S.H., David, S., van der Boog, A.T., Leemans, A., Verhoeff, J.J.: Changes in cortical thickness and volume after cranial radiation treatment: A systematic review. Radiotherapy and Oncology **135**, 33–42 (2019)
19. Puglisi, L., Alexander, D.C., Ravì, D.: Enhancing spatiotemporal disease progression models via latent diffusion and prior knowledge. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 173–183. Springer (2024)
20. Sakaguchi, Y., Kidokoro, H., Ogawa, C., Okai, Y., Ito, Y., Yamamoto, H., Ohno, A., Nakata, T., Tsuji, T., Nakane, T., et al.: Longitudinal findings of mri and pet in west syndrome with subtle focal cortical dysplasia. American Journal of Neuroradiology **39**(10), 1932–1937 (2018)
21. Sania, A., Pini, N., Nelson, M.E., Myers, M.M., Shuffrey, L.C., Lucchini, M., Elliott, A.J., Odendaal, H.J., Fifer, W.P.: K-nearest neighbor algorithm for imputing missing longitudinal prenatal alcohol data. Advances in Drug and Alcohol Research **4**, 13449 (2025)
22. Sotiras, A., Davatzikos, C., Paragios, N.: Deformable medical image registration: A survey. IEEE transactions on medical imaging **32**(7), 1153–1190 (2013)
23. Tiffon, C.: The impact of nutrition and environmental epigenetics on human health and disease. International journal of molecular sciences **19**(11), 3425 (2018)

24. van Timmeren, J., Bussink, J., Koopmans, P., Smeenk, R., Monshouwer, R.: Longitudinal image data for outcome modeling. Clinical Oncology (2024)
25. Tournier, J.D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.H., Connelly, A.: Mrtrix3: A fast, flexible and open software framework for medical image processing and visualisation. Neuroimage **202**, 116137 (2019)
26. Tustison, N.J., Avants, B.B., Gee, J.C.: Learning image-based spatial transformations via convolutional neural networks: A review. Magnetic resonance imaging **64**, 142–153 (2019)
27. Wang, J., Zhang, M.: Geo-sic: learning deformable geometric shapes in deep image classifiers. Advances in Neural Information Processing Systems **35**, 27994–28007 (2022)
28. Wasserthal, J., Breit, H.C., Meyer, M.T., Pradella, M., Hinck, D., Sauter, A.W., Heye, T., Boll, D.T., Cyriac, J., Yang, S., et al.: Totalsegmentator: robust segmentation of 104 anatomic structures in ct images. Radiology: Artificial Intelligence **5**(5), e230024 (2023)
29. Yoon, J.S., Zhang, C., Suk, H.I., Guo, J., Li, X.: Sadm: Sequence-aware diffusion model for longitudinal medical image generation. In: International Conference on Information Processing in Medical Imaging. pp. 388–400. Springer (2023)
30. Zhang, M., Singh, N., Fletcher, P.T.: Bayesian estimation of regularization and atlas building in diffeomorphic image registration. In: International conference on information processing in medical imaging. pp. 37–48. Springer (2013)
31. Zhang, Y., Shi, F., Wu, G., Wang, L., Yap, P.T., Shen, D.: Consistent spatial-temporal longitudinal atlas construction for developing infant brains. IEEE transactions on medical imaging **35**(12), 2568–2577 (2016)