

Learning Prototypes for Unsupervised Joint Generation of Conditional Templates and Atlases

Jichang Zhang¹, Tongtong Che¹, Feng Li¹, Lin Zhang², Chengbo Wang²,
Xiuying Wang², and Shuyu Li^{1*}

¹ State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, 100875, China

² School of Computer Science, The University of Sydney, Sydney, NSW 2006, Australia
shuyuli@bnu.edu.cn

Abstract. Conditional brain templates and atlases are essential for cohort-specific groupwise registration and for monitoring age-related or disease-imposed changes by quantifying anatomical and regional differences. However, most existing methods construct templates and atlases independently, leading to suboptimal quality of both. Besides, atlas construction often requires individual parcellation labelmaps, which are rarely available in practice. To address these limitations, we propose a unified framework that jointly generates conditional templates and atlases, requiring only an arbitrary pair of publicly available template and atlas. The model learns conditional prototypes in a shared latent space and maps them to templates and atlases via task-specific decoders, enabling effective interaction between anatomical and regional features. To further overcome the absence of individual labelmaps, we adopt a one-shot strategy that propagates the public atlas to individual subjects to build pseudo labelmaps, supervising the parcellation decoder. By optimizing prototypes across both latent and volumetric spaces over the entire population, the decoded templates maintain anatomical fidelity comparable to real MRIs, while both templates and atlases embody cohort-specific characteristics and model smooth variations across conditions. Comprehensive evaluations across four public datasets using eight multifaceted metrics demonstrate that our framework achieves favorable overall performance compared to recent state-of-the-art methods.

Keywords: Conditional Templates and Atlases · Prototype Generation · Brain MRI images · Registration.

1 Introduction

In neuroimaging analysis, brain templates and atlases are indispensable tools covering several uses [2,21]. A template is a canonical anatomical grayscale image used for spatial normalization of a group of MRI images via registration, enabling accurate comparison of the same brain region across a population [25,28].

* Corresponding author.

An atlas is a labelmap defined on the template that specifies the parcellation scheme and delineates distinct brain regions [30,10]. Numerous template and atlas pairs are available from public repositories [25,7,20]. However, these resources are typically created based on specific populations and cannot represent brain variability across diverse cohorts differing in age, disease status, or other conditions, often introducing systematic bias due to challenging registration [12].

A natural remedy is to learn conditional templates and atlases that form a series capturing condition-specific anatomical and regional characteristics. Ideally, conditional templates and atlases are intrinsically coupled through shared underlying structures and semantics: they should exhibit precise spatial correspondence, remain unbiased toward target cohorts, and faithfully model the trajectories of anatomical and regional changes that evolve across conditions. In practice, however, they are constructed separately, with atlas estimation performed merely as a post-hoc step after template construction [8,11,15].

Conditional templates are traditionally obtained by partitioning the dataset according to conditions and iteratively registering and averaging within each subset [3,17,6], a process that is prohibitively time-consuming. Modern deep-learning approaches achieve speed-ups and can be categorized into two paradigms. The first paradigm formulates the problem as conditional generation, mapping conditions to templates and modeling inter-condition relationships for smooth anatomical transitions [8,11,29]. Another paradigm, representation learning, operates in a group-wise registration fashion within a latent space by averaging latent representations to reconstruct high-quality templates, which preserves topology and ensures the resulting templates lie on an anatomically meaningful manifold [15,23,1]. Despite these strengths, both paradigms optimize purely in grayscale image space and miss explicit semantic information, such as the regional boundaries that could provide guidance to sharpen anatomical details of templates.

When conditional templates are inferred, one can employ automated labeling tools [11] or fuse multiple individual labelmaps [8] to produce corresponding atlases. However, automated tools and individual labelmaps that follow the desired parcellation scheme are rarely available, since they require a large amount of manual annotations by domain experts. In practice, the parcellation scheme is typically distributed to users in the form of a generic template and atlas pair. Thus, in unsupervised scenarios, the most feasible method is to warp the public atlas using deformation fields estimated by registering a public template to the conditional templates [29]. Yet, this approach is entirely intensity-driven and lacks region perception ability, thus alignment accuracy is severely limited in regions of interest [18]. Besides, potential deformation distortion can introduce topological artifacts such as holes and jagged edges into conditional atlases [9].

To address these challenges, we propose an unsupervised framework, named LatentAtlas, for the joint generation of conditional templates and atlases, requiring only an arbitrary public template and atlas pair as a prior. The key insight is to reformulate the construction of conditional templates and atlases as learning conditional prototypes in a unified latent space, where template reconstruction, atlas parcellation, and deformation estimation are jointly optimized

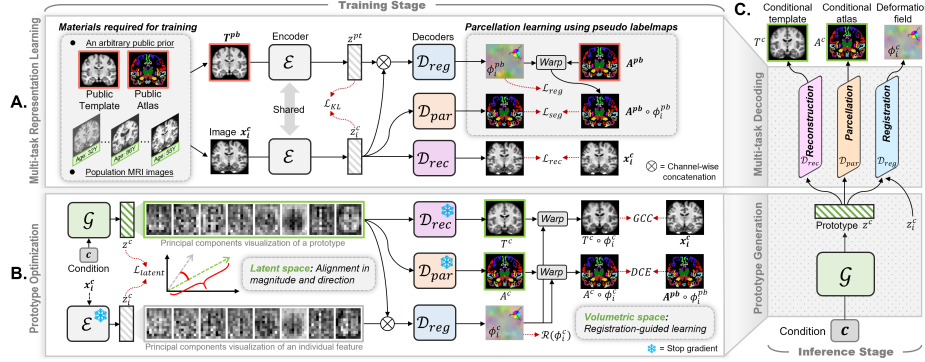


Fig. 1. The training stage consists of two steps: (A) a multi-decoder VAE is trained to encode images into latent space and map back to volumetric space comprising images, labelmaps, and deformation fields; and (B) a generator produces prototypes conditioned on inputs (e.g., age), which are encouraged to represent the specific cohorts in both latent and volumetric spaces. At the inference stage (C), a generated prototype can be jointly decoded into a template and an atlas given only the target condition.

via task-specific decoders. The learned prototype is a canonical latent representation encoding both anatomical and regional characteristics of the cohort, while the three tasks mutually reinforce each other through this shared factor [14,33]. In addition, the registration decoder exploits abundant unlabeled individual data by estimating deformation fields that warp the public atlas to each subject, generating pseudo labelmaps that supervise the parcellation decoder. This strategy leverages population-derived regularization to mitigate registration errors and topological artifacts, ensuring anatomically plausible atlas generation [31,13,33]. Within a single training process, we design a two-step procedure: first establishing the latent distribution and task-specific decoders on real MRIs, then optimizing the conditional generator to produce cohort-specific prototypes.

Experimentally, we adopt age, the primary demographic variable in neuroimaging studies, as the condition. We conduct comprehensive evaluation across four 3D MRI datasets, two public atlas priors, and eight metrics. The results demonstrate that LatentAtlas achieves superior performance, including template fidelity, atlas topological properties, centrality, and better age-trend fitting.

2 Method

2.1 Background: Conditional Template and Atlas Construction

Given a 3D MRI dataset $\mathcal{X} = \{\mathbf{x}_i^c\}_{i=1}^N$ where each image $\mathbf{x}_i^c$ is associated with a specific condition c (e.g., normalized age $\in [0, 1]$), constructing a conditional template involves generating an intensity image T^c and learning deformations $\{\phi_i^c : T^c \rightarrow \mathbf{x}_i^c\}$ constrained to be small and smooth, where ϕ_i^c is predicted by a registration network. Thus, the objective is:

$$T^c, \{\phi_i^c\} = \arg \min \mathbb{E}_{\mathbf{x}_i^c} [\mathcal{L}_{\text{sim}}(T^c \circ \phi_i^c, \mathbf{x}_i^c) + \mathcal{R}(\phi_i^c)], \text{ s.t. } \phi_i^c \in \Psi, \quad (1)$$

where $\circ$ denotes warping operation, $\mathcal{L}_{\text{sim}}$ measures dissimilarity, and $\mathcal{R}$ is a regularization term. $\phi \in \Psi$ is a topology-preserving diffeomorphism [3], defined by integrating a velocity field v via the ordinary differential equation $\frac{\partial \phi^{(t)}}{\partial t} = v(\phi^{(t)})$, where $\phi^{(0)} = \text{Id}$ is the identity transformation and t denotes time.

Once the conditional template is constructed, the corresponding conditional atlas A^c is obtained through label propagation via spatial transformation $\phi_c^{pb} : \mathbf{T}^{pb} \rightarrow T^c$ based on public templates $\mathbf{T}^{pb}$ and atlases $\mathbf{A}^{pb}$, as $A^c = \mathbf{A}^{pb} \circ \phi_c^{pb}$. Benefiting from the growing open-source practices, $(\mathbf{T}^{pb}, \mathbf{A}^{pb})$ are commonly released alongside publications on public repositories.

2.2 LatentAtlas

Independent generation of templates and atlases disregards their intrinsic coupling, which may produce geometrically inconsistent and semantically misaligned results. Moreover, most practical settings lack individual labelmaps, precluding supervised approaches. To address these challenges, we propose **LatentAtlas**, a novel framework that leverages public priors and population data for the joint generation of conditional templates and atlases. Fig. 1 illustrates the training and inference stages. Specifically, a multi-decoder Variational Autoencoder (VAE) is trained on MRI images and $(\mathbf{T}^{pb}, \mathbf{A}^{pb})$ to learn a unified multi-task latent space mappable to images, labelmaps, and deformation fields. A generator then estimates conditional latent features, which are regularized to align individual features within the same cohort while ensuring that the volumetric outputs satisfy registration-guided cohort centrality. At inference, the trained generator and decoders yield a conditional template and atlas from the input condition alone.

VAE backbone for multi-task representation learning. We propose a multi-decoder VAE to encode the brain MRI scans into latent features and decode them to images, labelmaps, and deformation fields.

Latent encoding. An encoder $\mathcal{E}_{\theta_e}$ maps $\mathbf{x}_i^c$ and $\mathbf{T}^{pb}$ to latent features z_i^c and z^{pb} by predicting distributional parameters $(\mu, \log \sigma^2)$ and sampling from $\mathcal{N}(\mu, \sigma^2)$. We employ gated normalization layers to handle the intensity distribution mismatch between $\mathbf{x}_i^c$ and $\mathbf{T}^{pb}$, inspired by practices in multi-modal learning [23].

Task-specific decoding. The encoded and normalized features z_i^c and z^{pb} are fed into decoders of reconstruction $\mathcal{D}_{\theta_{rec}}$, parcellation $\mathcal{D}_{\theta_{par}}$, and registration $\mathcal{D}_{\theta_{reg}}$, with the forward pass illustrated in Fig. 1A. A key challenge is the absence of individual labelmaps for optimizing the parcellation decoder. To overcome this, we adopt a one-shot strategy, i.e., applying ϕ_i^{pb} decoded from $\mathcal{D}_{\theta_{reg}}[z^{pb} \otimes z_i^c]$ to warping $\mathbf{A}^{pb}$ that serve as pseudo labelmaps for supervising $\hat{y}_i^c = \mathcal{D}_{\theta_{par}}[z_i^c]$. This strategy exploits registration-preserved structural correspondence, which provides effective supervision despite potential label noises [31]. The parameters $\theta_e, \theta_{rec}, \theta_{reg}, \theta_{par}$ are jointly optimized via the multi-task loss $\mathcal{L}_{mtl}$, as:

$$\mathcal{L}_{mtl} = \lambda_1 \mathcal{L}_{rec} + \lambda_2 \mathcal{L}_{par} + \lambda_3 \mathcal{L}_{reg} + \lambda_{KL} \mathcal{L}_{KL}, \quad (2)$$

where $\mathcal{L}_{rec} = MSE(\mathcal{D}_{\theta_{rec}}[z_i^c], \mathbf{x}_i^c)$, $\mathcal{L}_{par} = DCE(\hat{y}_i^c, \mathbf{A}^{pb} \circ \phi_i^{pb})$, and $\mathcal{L}_{reg} = GCC(\mathbf{T}^{pb} \circ \phi_i^{pb}, \mathbf{x}_i^c) + BE(\phi_i^{pb})$, and $\mathcal{L}_{KL}$ is the KL divergence that regularizes the latent distribution toward a standard normal prior. Here, MSE , DCE , GCC , and BE denote mean squared error, Dice and cross-entropy losses, global cross-correlation, and bending energy, respectively.

Conditional prototype generation. Motivated by prototype learning [22,32], in which a prototype represents the centroid of a sample cluster or semantic class in latent space for feature matching in downstream tasks, we aim to learn conditional prototypes that capture the canonical anatomical and regional representations of condition-specific cohorts. Rather than separately estimating T^c and A^c , jointly decoding prototypes into both capitalizes on the inherent inter-dependencies between anatomical and regional information.

The generation of conditional prototypes can be formulated as $z^c = \mathcal{G}_{\theta_g}[\mathbf{c}]$. To better capture inter-cohort variations, we follow [11] and design $\mathcal{G}_{\theta_g}$ as a convolutional network equipped with feature-wise linear modulation (FiLM [27]) layers.

Optimization of prototypes. The multi-decoder VAE backbone learns a latent manifold underlying real MRI data while providing multi-task decoding capabilities. In conjunction with this backbone, we design optimization objectives for $\mathcal{G}_{\theta_g}$ in both the latent space and the volumetric space.

In the latent space, the conditional prototype should lie in close proximity to the region of the learned manifold occupied by individual features sharing the same condition, thereby capturing the representative characteristics of the cohort. To enforce this proximity, we jointly constrain the prototype z^c and individual features $\{z_i^c\}_{i=1}^N$ through alignment in both magnitude and direction, using $L2$ and cosine similarity losses:

$$\mathcal{L}_{latent} = \beta_1 \|z^c - z_i^c\|_2^2 + \beta_2 \left(1 - \frac{z^c \cdot z_i^c}{\|z^c\|_2 \|z_i^c\|_2}\right). \quad (3)$$

In the volumetric space, we enforce similarity between the warped conditional template $T^c \circ \phi_i^c$ and $\mathbf{x}_i^c$ to ensure that the decoded T^c captures appearance and shape consistent with the individual anatomy. To promote accurate conditional atlases, the warped atlas $A^c \circ \phi_i^c$ is aligned with the individual pseudo labelmaps. To preserve cohort centrality, we apply the regularization $\mathcal{R}(\phi_i^c) = \|\phi_i^c\|_2^2 + \|\bar{\phi}_i^c\|_2^2 + BE(\phi_i^c)$, where $\bar{\phi}_i^c$ denotes the cohort-level moving average deformation.

Gradient flow management. The complexity of our framework requires careful gradient flow design. First, prototype learning benefits from individual data, but not vice versa; thus, we detach $\mathcal{L}_{latent}$ from $\mathcal{E}_{\theta_e}$, so that it updates only $\mathcal{G}_{\theta_g}$. Second, gradients from prototype-related losses are detached from $\mathcal{D}_{\theta_{rec}}$ and $\mathcal{D}_{\theta_{par}}$, as these gradient signals exhibit limited contribution to their training. Lastly, the centrality regularization $\mathcal{R}(\phi_i^c)$ is prevented from updating $\mathcal{D}_{\theta_{reg}}$, ensuring it focuses on anatomical alignment rather than population statistics.

Table 1. Intrinsic qualities of the constructed templates and atlases. Lower *T. Sharp.* indicates sharper templates; lower *Trans.*, smoother transitions; higher (less negative) *A. Topo.*, simpler topology with fewer holes. Best marked with , second with .

Metric	MIND Atlas				IBSR Atlas			
	AtlasMorph	DiffDef	InGroup	Ours	AtlasMorph	DiffDef	InGroup	Ours
T. Sharp.↓	.92±.02	.91±.01	.91±.01	.90±.01	.92±.02	.97±.01	.91±.01	.90±.01
T. Trans.↓	.17±.11	.74±.90	8.7±3.4	.33±.18	.17±.11	1.1±1.8	8.7±3.4	.29±.10
A. Trans.↓	4.2±1.4	10±5.9	29±5.4	3.6±0.8	3.3±1.2	9.2±7.2	28±4.0	3.1±.84
A. Topo.↑	-3260±44	-3140±83	-3378±75	-1656±45	-44±9	-39±25	-46±17	-27±3

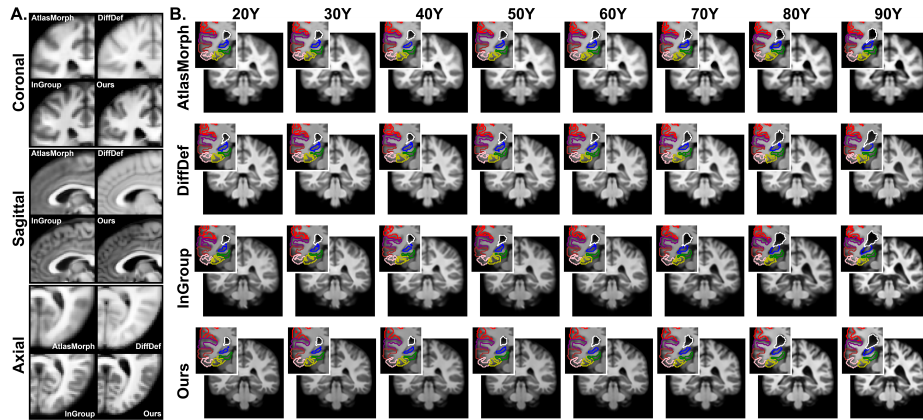


Fig. 2. **A.** Local magnified views of the IBSR-based 55Y templates in the three anatomical planes. **B.** Conditional templates and atlases (local view) based on the MIND atlas.

3 Experimental Evaluation

Materials: Data and Baselines. We evaluate our framework on four publicly available 3D T1-MRI datasets comprising 831 scans in total. We train on OASIS [24] (314 subjects, 18-96Y) and test on three datasets: IXI [16] (400 subjects, 20-90Y), MIND [19] (99 subjects with annotated labelmaps, 19-61Y, with a template and a 95-class atlas), and IBSR [18] (18 subjects with annotated labelmaps, youth-51Y, with a MNI152 template [25] and a 34-class atlas derived using MALF [4]). All subjects are healthy, and all images are affinely registered to the MNI152 template, skull-stripped, normalized, and downsampled to 2 mm³ isotropic resolution with volume size 80 × 96 × 80.

We compare LatentAtlas against three state-of-the-art methods: AtlasMorph [8], DiffDef [29], and InGroup [15]. The generator of AtlasMorph is adapted to include FiLM layers. InGroup constructs conditional templates using age-nearest neighbors from the training set. AtlasMorph and InGroup propagate the MIND or IBSR atlas onto their generated templates using a pre-trained registration

Table 2. Comparison of different methods across four datasets and four metrics. ♦: training set; ◇: test set. ↑: higher is better; ↓: lower is better. *: statistical significance with $p < 0.01$ in paired t-test.

Methods	OASIS♦			IXI◇			MIND◇		IBSR◇	
	Mov ↓	Norm ↓	LPIPS ↓	Mov ↓	Norm ↓	LPIPS ↓	Norm ↓	DSC ↑	Norm ↓	DSC ↑
AtlasMorph	329±42	739±93	.38±.02	488±62	762±88	.43±.03	811±74	.50±.02	709±63	.65±.09
DiffDef _(IBSR)	624±43	871±65	.36±.02	564±31	789±70	.33±.03	N/A	N/A	831±99	.61±.09
DiffDef _(MIND)	310±37	722±89	.25±.02	502±53	769±85	.31±.05	837±79	.52±.02	N/A	N/A
InGroup	236±39*	692±96	.21±.02	484±51	756±82	.33±.05	804±78	.49±.02	729±68	.65±.02
Ours _(IBSR)	246±37	667±98	.21±.02	456±74*	722±81*	.31±.05	N/A	N/A	702±67	.66±.02
Ours _(MIND)	251±40	660±96*	.20±.02	463±48	735±81	.30±.04*	759±71*	.52±.03	N/A	N/A

network, whereas DiffDef employs learned conditional deformation fields. All comparative methods use their originally reported hyperparameters.

Architectural and Training Details. Our LatentAtlas is implemented using PyTorch and MONAI [5]. For multi-task learning, we build the encoder-decoder architectures from MONAI’s AutoencoderKL module with latent shape $(B, 64, 10, 12, 10)$. The registration decoder incorporates an additional convolutional network for concatenated feature inputs. Conditional generator is a convolutional network with FiLM layers. To facilitate learning, we initialize the encoder and reconstruction decoder using a pre-trained reconstruction VAE on the training set, which DiffDef also utilizes. We set the loss weights $[\lambda_{KL}, \beta_1] = [10^{-8}, 0.1]$, while all other weights default to 1. The weights of the three terms in $\mathcal{R}$ are $[1, 0.01, 1]$. Training uses Adam optimizer with learning rate 10^{-4} , batch size 1, for 300 epochs on an A6000 GPU. Our source code is publicly available at <https://github.com/LSYLAB/LatentAtlas>.

Evaluation Metrics. We evaluate intrinsic qualities including template sharpness (EFC), transition smoothness between adjacent templates or atlases (SSIM-Loss), and atlas topology (Euler number). We also measure template centrality via deformation cost to training cohorts and generalizability to test data. This includes the norm of moving average of deformations (Mov) and the average deformation norms to target images (Norms), defined as $\|\bar{\phi}_i^c\|_2^2$ and $\|\phi_i^c\|_2^2$, averaged within the age-specific cohort. Besides, we compute the Perceptual score (LPIPS [34]) of generated templates and real data to evaluate the fidelity. Yet optimal centrality necessitates a trade-off between minimizing registration cost and preserving registration accuracy. Therefore, we further measure the average Dice coefficients (DSC) for individual labelmaps deformed from the template and atlas to the target on MIND and IBSR.

Results. Table 1 presents intrinsic quantitative evaluations of the constructed conditional templates and atlases. Fig. 2 visualizes conditional templates and

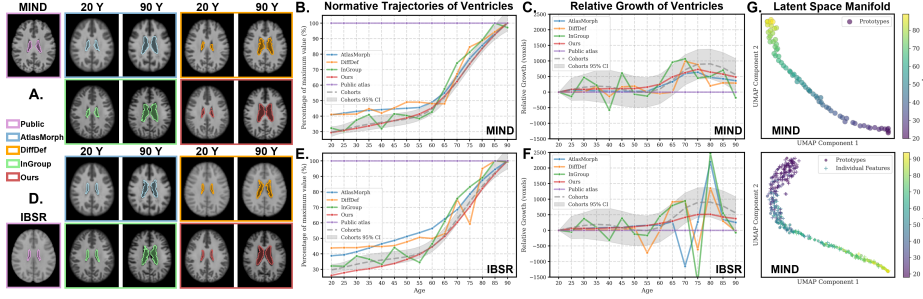


Fig. 3. **A.** MIND-based conditional templates with ventricles defined from conditional atlases (ages of 20 and 90). **B.** Normative trajectories of ventricles. **C.** Relative growth of ventricles (age-to-age residuals). **D-F.** IBSR-based counterparts. **G.** UMAP visualization of prototypes (top) and joint embedding with individual features (bottom).

atlases constructed using IBSR and MIND atlas priors. While the EFC-based sharpness metrics are comparable across methods, the visual comparison further highlights that our method preserves richer anatomical structures and clearer local details. For age-wise transitions, AtlasMorph achieves smoother template transitions, whereas our method shows more favorable atlas-level transitions. In contrast, InGroup exhibits poor cross-age continuity due to its reliance on cohort-specific inference, resulting in abrupt shifts between adjacent templates (see Fig. 2).

However, smooth transitions may reflect either genuine age progression or over-smoothing. We therefore assess template centrality to confirm preservation of age-specific characteristics. Table 2 shows that our method achieves comparable centrality and fidelity to InGroup on source data. On out-of-distribution samples from two external cohorts with publicly available gold-standard annotations, our method achieves comparable DSC with lower deformation cost and better image fidelity, suggesting it captures age-specific features rather than overfitting to individual variability within the training set. Notably, these improvements in transition smoothness and centrality are achieved while yielding relatively simpler atlas topology compared with competing methods under the same prior (A. Topo. in Table 1). All methods exhibit negligible deformation folding (around 0.03%).

Fig. 3B and E show age-related ventricular volume trajectories derived from the constructed conditional atlases, a structure that expands significantly with aging. Using either MIND or IBSR priors, our method produces trajectories that most closely match the OASIS observations, without the irregular fluctuations observed in DiffDef and InGroup due to template bias and direct registration inaccuracies. Importantly, our method yields consistent trajectories irrespective of the public prior used, demonstrating its robustness. Finally, we apply UMAP [26] to interpret the learned prototypes. As shown in Fig. 3G, the latent manifold reveals that the prototypes form a smooth, age-ordered continuum while remaining representative of the individual feature distributions.

4 Conclusion

In conclusion, we propose LatentAtlas, an unsupervised framework that learns conditional prototypes to jointly generate conditional brain templates and atlases. At its core, LatentAtlas unifies template and atlas generation within a shared latent space, allowing their learning to mutually reinforce one another. Further, the pseudo labeling strategy uses abundant unlabeled data to drive the LatentAtlas training from merely a single public template and atlas pair, boosting framework performance and practical applicability. Rigorous multi-metric evaluation demonstrates the superior quality of the conditional brain templates and atlases generated by LatentAtlas. Future work will focus on extending LatentAtlas to additional covariates beyond age and validating its applicability across more imaging modalities and larger-scale datasets.

Acknowledgments. This study was funded by National Natural Science Foundation of China [Nos. 32271146 and 81622025], and the Startup Funds for Talents at Beijing Normal University, China.

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

References

1. Abulnaga, S.M., Hoopes, A., Dey, N., Hoffmann, M., Fischl, B., Gutttag, J., Dalca, A.: Multimorph: On-demand atlas construction. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 30906–30917 (2025)
2. Alexander-Bloch, A., Giedd, J.N., Bullmore, E.: Imaging structural co-variance between human brain regions. *Nature reviews neuroscience* **14**(5), 322–336 (2013)
3. Avants, B., Gee, J.C.: Geodesic estimation for large deformation anatomical shape averaging and interpolation. *Neuroimage* **23**, S139–S150 (2004)
4. Avants, B.B., Tustison, N., Song, G., et al.: Advanced normalization tools (ants). *Insight j* **2**(365), 1–35 (2009)
5. Cardoso, M.J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., Murrey, B., Myronenko, A., Zhao, C., Yang, D., et al.: Monai: An open-source framework for deep learning in healthcare. *arXiv preprint arXiv:2211.02701* (2022)
6. Che, T., Zhang, L., Zeng, D., Zhao, Y., Bai, H., Zhang, J., Wang, X., Li, S.: Nested hierarchical group-wise registration with a graph-based subgrouping strategy for efficient template construction. *Medical Image Analysis* p. 103624 (2025)
7. Ciric, R., Thompson, W.H., Lorenz, R., Goncalves, M., MacNicol, E.E., Markiewicz, C.J., Halchenko, Y.O., Ghosh, S.S., Gorgolewski, K.J., Poldrack, R.A., et al.: Templateflow: Fair-sharing of multi-scale, multi-species brain models. *Nature methods* **19**(12), 1568–1571 (2022)
8. Dalca, A., Rakic, M., Gutttag, J., Sabuncu, M.: Learning conditional deformable templates with convolutional networks. *Advances in neural information processing systems* **32** (2019)
9. Dalca, A.V., Balakrishnan, G., Gutttag, J., Sabuncu, M.R.: Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces. *Medical image analysis* **57**, 226–236 (2019)

10. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., et al.: An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *Neuroimage* **31**(3), 968–980 (2006)
11. Dey, N., Ren, M., Dalca, A.V., Gerig, G.: Generative adversarial registration for improved conditional deformable templates. In: *Proceedings of the IEEE/CVF international conference on computer vision*. pp. 3929–3941 (2021)
12. Evans, A.C., Janke, A.L., Collins, D.L., Baillet, S.: Brain templates and atlases. *Neuroimage* **62**(2), 911–922 (2012)
13. He, Y., Ge, R., Qi, X., Chen, Y., Wu, J., Coatrieux, J.L., Yang, G., Li, S.: Learning better registration to learn better few-shot medical image segmentation: Authenticity, diversity, and robustness. *IEEE Transactions on Neural Networks and Learning Systems* **35**(2), 2588–2601 (2022)
14. He, Y., Yang, G., Ge, R., Chen, Y., Coatrieux, J.L., Wang, B., Li, S.: Geometric visual similarity learning in 3d medical image self-supervised pre-training. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 9538–9547 (2023)
15. He, Z., Chung, A.C.: Instantgroup: Instant template generation for scalable group of brain mri registration. *IEEE Transactions on Image Processing* (2025)
16. IXI Dataset: IXI dataset (2018). Available online at: <https://brain-development.org/ixi-dataset/> (2018)
17. Joshi, S., Davis, B., Jomier, M., Gerig, G.: Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage* **23**, S151–S160 (2004)
18. Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.C., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., et al.: Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration. *Neuroimage* **46**(3), 786–802 (2009)
19. Klein, A., Tourville, J.: 101 labeled brain images and a consistent human cortical labeling protocol. *Frontiers in neuroscience* **6**, 171 (2012)
20. Lawrence, R.M., Bridgeford, E.W., Myers, P.E., Arvapalli, G.C., Ramachandran, S.C., Pisner, D.A., Frank, P.F., Lemmer, A.D., Nikolaidis, A., Vogelstein, J.T.: Standardizing human brain parcellations. *Scientific data* **8**(1), 78 (2021)
21. Leergaard, T.B., Bjaalie, J.G.: Atlas-based data integration for mapping the connections and architecture of the brain. *Science* **378**(6619), 488–492 (2022)
22. Li, O., Liu, H., Chen, C., Rudin, C.: Deep learning for case-based reasoning through prototypes: A neural network that explains its predictions. In: *Proceedings of the AAAI conference on artificial intelligence*. vol. 32 (2018)
23. Luo, X., Wang, X., Shapiro, L., Yuan, C., Feng, J., Zhuang, X.: Bayesian unsupervised disentanglement of anatomy and geometry for deep groupwise image registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2025)
24. Marcus, D.S., Wang, T.H., Parker, J., Csernansky, J.G., Morris, J.C., Buckner, R.L.: Oasis: cross-sectional mri data in young, middle aged, nondemented, and demented older adults. *Journal of cognitive neuroscience* **19**(9), 1498–1507 (2007)
25. Mazziotta, J.C., Toga, A.W., Evans, A., Fox, P., Lancaster, J., et al.: A probabilistic atlas of the human brain: theory and rationale for its development. *Neuroimage* **2**(2), 89–101 (1995)
26. McInnes, L., Healy, J., Melville, J.: Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426* (2018)
27. Perez, E., Strub, F., De Vries, H., Dumoulin, V., Courville, A.: Film: Visual reasoning with a general conditioning layer. In: *Proceedings of the AAAI conference on artificial intelligence*. vol. 32 (2018)

28. Rohlfing, T., Zahr, N.M., Sullivan, E.V., Pfefferbaum, A.: The sri24 multichannel atlas of normal adult human brain structure. *Human brain mapping* **31**(5), 798–819 (2010)
29. Starck, S., Sideri-Lampretsa, V., Kainz, B., Menten, M.J., Mueller, T.T., Rueckert, D.: Diff-def: Diffusion-generated deformation fields for conditional atlases. *IEEE Transactions on Medical Imaging* (2025)
30. Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., 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)
31. Xu, Z., Niethammer, M.: Deepatlas: Joint semi-supervised learning of image registration and segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 420–429. Springer (2019)
32. Yuan, Y., Wang, X., Yang, X., Heng, P.A.: Effective semi-supervised medical image segmentation with probabilistic representations and prototype learning. *IEEE Transactions on Medical Imaging* **44**(3), 1181–1193 (2024)
33. Zhang, L., Ning, G., Liang, H., Han, B., Liao, H.: One-shot neuroanatomy segmentation through online data augmentation and confidence aware pseudo label. *Medical Image Analysis* **95**, 103182 (2024)
34. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 586–595 (2018)