IMPERIAL

Multi-Modality Groupwise Medical Image Registration: A Generative Perspective

Xinzhe Luo
July 11, 2024

Motivation

Background

Multi-Modality Image Registration

A preprocessing step to relate, compare, and combine information from multi-modal medical imaging for the sake of image-based diagnosis.

An example from cardiac imaging

$\mathcal{X}$-Metric: Information-Theoretic Groupwise Registration

Generalizing Mutual Informaiton to High Dimensionality

For a fixed image $F : \Omega_F \to \mathbb{R}$ and a moving image $M : \Omega_M \to \mathbb{R}$, the transformation $\widehat{\phi} : \Omega_F \to \Omega_M$ that registers them is defined by

$$\widehat{\phi} = \arg\max_{\phi} \left\{ I(F, M \circ \phi) + R(\phi) \right\}, \tag{1}$$

where $I(X, Y) = H(X) - H(X \mid Y) = D_{KL}\left[P(X, Y) \parallel P(X)P(Y)\right]$ is the mutual information (MI) between the two images.

Generalizing MI to high dimensionality is non-trivial due to the curse of dimensionality, since the kernel density estimator for estimating $P(\mathbf{U})$ is intractable for large image groups. One such generalization is the total correlation. For N images $\mathbf{U} = \{U_j\}_{j=1}^{N}$ and transformations $\phi \triangleq \{\phi_j\}_{j=1}^{N}$, it is defined by

$$C(\mathbf{U}[\phi]) \triangleq D_{KL}\left[P(\mathbf{U}[\phi]) \parallel \prod_{j=1}^{N} P(U_j \circ \phi_j)\right] = \sum_{j=1}^{N} H(U_j \circ \phi_j) - H(\mathbf{U}[\phi]), \tag{2}$$

where $H(\cdot)$ is the Shannon entropy, and $\mathbf{U}[\phi] \triangleq \{U_j \circ \phi_j\}_{j=1}^{N}$.

$\mathcal{X}$-Metric: Information-Theoretic Groupwise Registration

Generalizing Mutual Informaiton to High Dimensionality

For a fixed image $F : \Omega_F \to \mathbb{R}$ and a moving image $M : \Omega_M \to \mathbb{R}$, the transformation $\widehat{\phi} : \Omega_F \to \Omega_M$ that registers them is defined by

$$\widehat{\phi} = \arg\max_{\phi} \{I(F, M \circ \phi) + R(\phi)\}, \tag{1}$$

where $I(X, Y) = H(X) - H(X \mid Y) = D_{KL}\left[P(X, Y) \parallel P(X)P(Y)\right]$ is the mutual information (MI) between the two images.

Generalizing MI to high dimensionality is non-trivial due to the curse of dimensionality, since the kernel density estimator for estimating $P(\mathbf{U})$ is intractable for large image groups. One such generalization is the total correlation. For N images $\mathbf{U} = \{U_j\}_{j=1}^{N}$ and transformations $\phi \triangleq \{\phi_j\}_{j=1}^{N}$, it is defined by

$$C(\mathbf{U}[\phi]) \triangleq D_{KL}\left[P(\mathbf{U}[\phi]) \parallel \prod_{j=1}^{N} P(U_j \circ \phi_j)\right] = \sum_{j=1}^{N} H(U_j \circ \phi_j) - H(\mathbf{U}[\phi]), \tag{2}$$

where $H(\cdot)$ is the Shannon entropy, and $\mathbf{U}[\phi] \triangleq \{U_j \circ \phi_j\}_{j=1}^{N}$.

$\mathcal{X}$-Metric: Information-Theoretic Groupwise Registration
A Generative Perspective

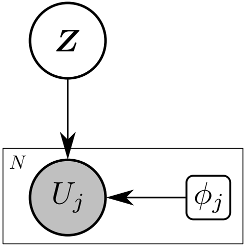
For the multi-modality images $\mathbf{U} = \{U_j\}_{j=1}^N$ to be registered, they are assumed to be generated by a common latent variable $\mathbf{Z}$ representing the common anatomy, i.e., $P(\mathbf{U}|\mathbf{Z}) = \prod_{j=1}^N P(U_j|\mathbf{Z})$.

The reduction of the uncertainty in $\mathbf{Z}$ due to observations $\mathbf{U}[\phi]$ can be measured by

$$I(\mathbf{U}[\phi], \mathbf{Z}) = H(\mathbf{Z}) - H(\mathbf{Z} \mid \mathbf{U}[\phi]) = H(\mathbf{U}[\phi]) - H(\mathbf{U}[\phi] \mid \mathbf{Z}). \quad (3)$$

Due to the conditional independence $P(\mathbf{U}[\phi] \mid \mathbf{Z}) = \prod_{j=1}^N P(U_j \circ \phi_j \mid \mathbf{Z})$, we have

$$I(\mathbf{U}[\phi], \mathbf{Z}) = H(\mathbf{U}[\phi]) - \sum_{j=1}^N H(U_j \circ \phi_j \mid \mathbf{Z}). \quad (4)$$



Graphical model.

$\mathcal{X}$-Metric: Information-Theoretic Groupwise Registration

A Generative Perspective

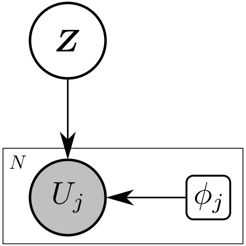
For the multi-modality images $\mathbf{U} = \{\mathrm{U_j}\}_{\mathrm{j}=1}^{\mathrm{N}}$ to be registered, they are assumed to be generated by a common latent variable $\mathbf{Z}$ representing the common anatomy, i.e., $\mathrm{P}(\mathbf{U}|\mathbf{Z}) = \prod_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{P}(\mathrm{U_j}|\mathbf{Z})$.

The reduction of the uncertainty in $\mathbf{Z}$ due to observations $\mathbf{U}[\phi]$ can be measured by

$$\mathrm{I}(\mathbf{U}[\phi], \mathbf{Z}) = \mathrm{H}(\mathbf{Z}) - \mathrm{H}(\mathbf{Z} \mid \mathbf{U}[\phi]) = \mathrm{H}(\mathbf{U}[\phi]) - \mathrm{H}(\mathbf{U}[\phi] \mid \mathbf{Z}). \tag{3}$$

Due to the conditional independence $\mathrm{P}(\mathbf{U}[\phi] \mid \mathbf{Z}) = \prod_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{P}(\mathrm{U_j} \circ \phi_{\mathrm{j}} \mid \mathbf{Z})$, we have

$$\mathrm{I}(\mathbf{U}[\phi], \mathbf{Z}) = \mathrm{H}(\mathbf{U}[\phi]) - \sum_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{H}(\mathrm{U_j} \circ \phi_{\mathrm{j}} \mid \mathbf{Z}). \tag{4}$$



Graphical model.

$\mathcal{X}$-Metric: Information-Theoretic Groupwise Registration

A Generative Perspective

- Both $\mathsf{C}(\mathbf{U})$ and $\mathsf{I}(\mathbf{U}, \mathbf{Z})$ can be used as a groupwise similarity measure. However, both measures involve the annoying joint entropy term $\mathsf{H}(\mathbf{U})$.
- The combination of the two measures happens to cancel out this term, leading to the proposed $\mathcal{X}$-metric:

$$\mathcal{X}(\mathbf{U}, \mathbf{Z}) \triangleq \mathsf{C}(\mathbf{U}) + \mathsf{I}(\mathbf{U}, \mathbf{Z}) = \sum_{\mathrm{j}=1}^{\mathrm{N}} \mathsf{I}(\mathrm{U_j}, \mathbf{Z}). \tag{5}$$

- The optimization problem for groupwise registration becomes

$$\widehat{\phi} = \arg\max_{\boldsymbol{\phi}} \max_{\boldsymbol{\alpha}} \mathcal{X}(\mathbf{U}[\phi], \mathbf{Z}), \tag{6}$$

where α are the nuisance parameters within the generative model.

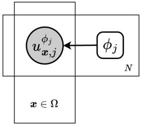
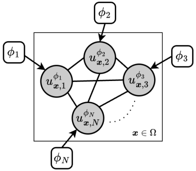
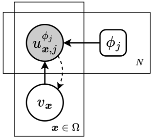
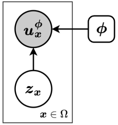
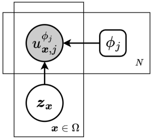
Algorithm 1. $\mathcal{X}$-CoReg

Data: The observed images $U = \{U_j\}_{j=1}^N$;
Input: Number of iterations T, regularization coefficient λ, registration step size η;
Output: The estimated spatial transformations $\widehat{\boldsymbol{\phi}} = \{\widehat{\phi}_j\}_{j=1}^N$;

```
Initialization: φ_j^[0] ≜ id for j = 1,...,N; initialize π^[0] and Γ^[0] by (22);
for t = 0,...,T−1 do
    /* Update the common-space parameters */
    γ_{x,k}^[t+1] ≜ π_k^[t] ∏_{j=1}^N f_jk^[t](μ_j; φ_j^[t]) / Σ_{k=1}^K π_k^[t] ∏_{j=1}^N f_jk^[t](μ_j; φ_j^[t]);
    π_k^[t+1] ≜ Σ_{x∈Ω^{φ[t]}} γ_{x,k}^[t+1] / Σ_{k=1}^K Σ_{x∈Ω^{φ[t]}} γ_{x,k}^[t+1];
    /* Update the spatial transformations */
    φ^[t+1] = φ^[t] − η · ∇L(φ | U; Γ^[t+1]) |_{φ=φ^[t]};
    if L converges then
        break loop;
return φ̂ = φ^[T].
```

$\mathcal{X}$-Metric: Information-Theoretic Groupwise Registration

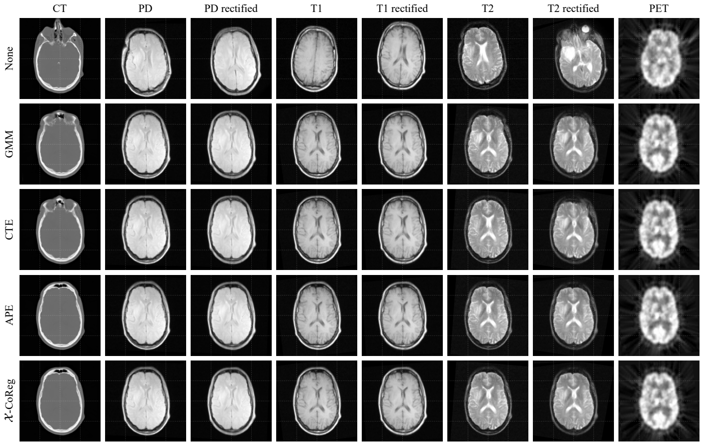
A Generative Perspective

Method	CG	APE	CTE	GMM	$\mathcal{X}$-CoReg
Graphical model					
Computational complexity	$\mathcal{O}(N^2\vert\Omega\vert)$	$\mathcal{O}(L^2N^2\vert\Omega\vert)$	$\mathcal{O}(N^2\vert\Omega\vert + N^3 + L^2N\vert\Omega\vert)$	$\mathcal{O}(KN^3\vert\Omega\vert)$	$\mathcal{O}(KLN\vert\Omega\vert)$

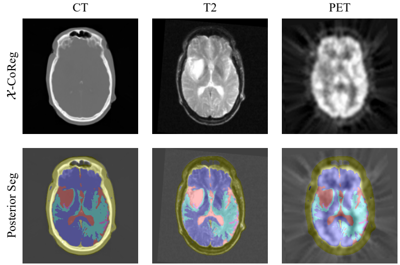
- **CG**: E. G. Learned-Miller, "Data driven image models through continuous joint alignment," TPAMI 2006.
- **APE**: C. Wachinger and N. Navab, "Simultaneous registration of multiple images: Similarity metrics and efficient optimization," TPAMI 2013.
- **CTE**: M. Polfliet et al., "Intrasubject multimodal groupwise registration with the conditional template entropy," MedIA 2018.
- **GMM**: J. Orchard and R. Mann, "Registering a multi-sensor ensemble of images," TIP 2010.

Experiments

Multi-Modality Groupwise Rigid Registration



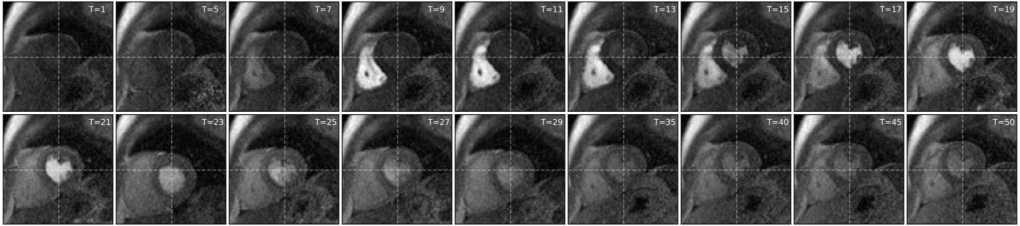
Method	gRE (mm)	Time-r	GPU-r	p-value
None	40.65 ± 4.29*	–	–	$< 10^{-10}$
GMM [16]	15.84 ± 7.78*	0.57	0.96	$< 10^{-10}$
CTE [6]	13.46 ± 5.04*	0.98	1.01	$< 10^{-10}$
APE [4]	5.36 ± 2.60	2.02	2.00	0.61
$\mathcal{X}$-CoReg	5.52 ± 1.90	1.00	1.00	–

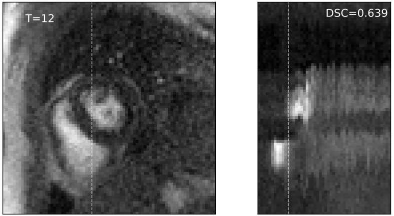


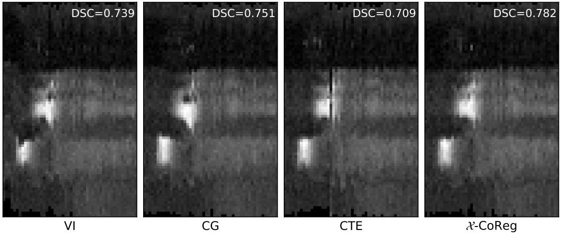
The posterior segmentation overlaid on the registered images; major anatomical structures are revealed from the posterior segmentation.

Experiments

Spatiotemporal Motion Correction on First-Pass Cardiac Perfusion MR







Before motion correction.

After motion correction.

Unsupervised Disentanglement of Anatomy and Geometry
Motivation

- $\mathcal{X}$-metric incorporates the generative perspective into the design of a new groupwise similarity metric. However, it still lacks scalability on large image groups.
- Given the common anatomy $\mathbf{Z}$, the observed image U_j is generated through the composition of an imaging functional f_j and a spatial transformation ϕ_j^{-1},

$$\mathsf{U}_\mathsf{j} = \mathsf{f}_\mathsf{j}(\mathbf{Z}) \circ \phi_\mathsf{j}^{-1}, \tag{7}$$

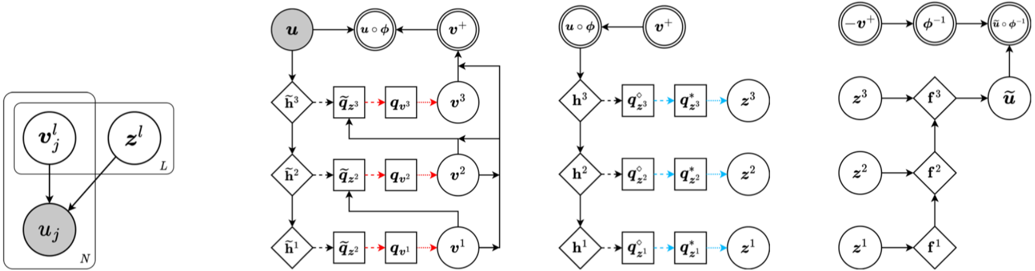
where f_j is assumed to be spatially equivariant w.r.t. the group of diffeomorphisms $\mathcal{G}$, i.e.

$$\mathsf{f}_\mathsf{j}(\mathbf{Z}) \circ \phi_\mathsf{j}^{-1} = \mathsf{f}_\mathsf{j}(\mathbf{Z} \circ \phi_\mathsf{j}^{-1}), \quad \forall\, \phi_\mathsf{j} \in \mathcal{G}. \tag{8}$$

- Given the imaging functional, the variation of U_j comes from the underlying generative factors $\mathbf{Z}$ and ϕ_j. From the perspective of disentanglement learning, we want a model to simultaneously infer them, which also leads to groupwise registration.

Unsupervised Disentanglement of Anatomy and Geometry

Hierarchical Bayesian Inference



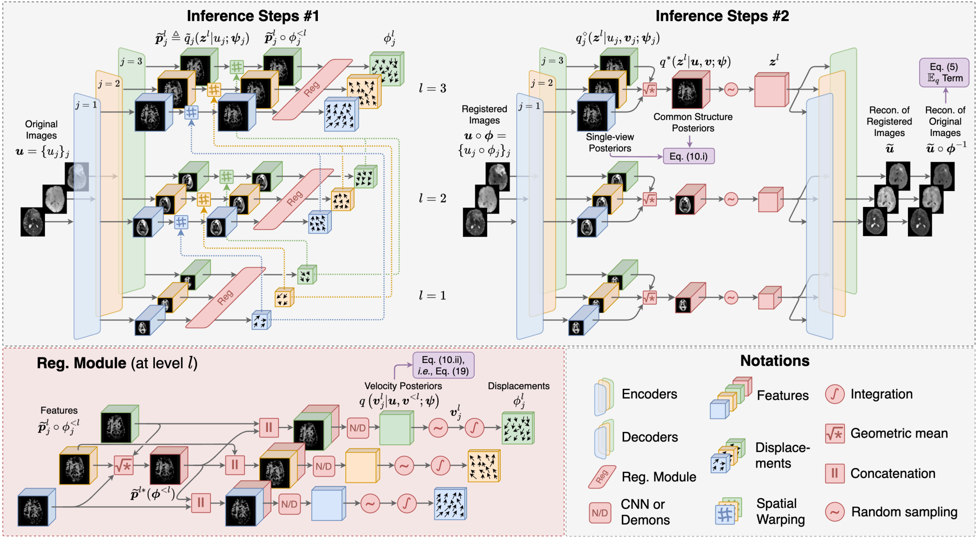
(a) Generative model. (b) Inference steps #1. (c) Inference steps #2. (d) Generation steps.

The objective function we optimize is the evidence lower bound (ELBO):

$$\mathcal{L}(\boldsymbol{\theta}, \psi \mid \mathbf{u}) \triangleq \mathbb{E}_{\mathrm{q}(\mathbf{z},\mathbf{v}\mid\mathbf{u};\psi)}[\log \mathrm{p}(\mathbf{u} \mid \mathbf{z}, \mathbf{v}; \boldsymbol{\theta})] - \mathrm{D}_{\mathrm{KL}}\left[\mathrm{q}(\mathbf{z}, \mathbf{v} \mid \mathbf{u}; \psi) \parallel \mathrm{p}(\mathbf{z})\mathrm{p}(\mathbf{v})\right], \tag{9}$$

$$\mathrm{D}_{\mathrm{KL}}\left[\mathrm{q}(\mathbf{z}, \mathbf{v} \mid \mathbf{u}; \psi) \parallel \mathrm{p}(\mathbf{z})\mathrm{p}(\mathbf{v})\right] = \mathbb{E}_{\mathrm{q}(\mathbf{v}\mid\mathbf{u};\psi)}\left[\mathrm{D}_{\mathrm{KL}}\left[\mathrm{q}(\mathbf{z} \mid \mathbf{u}, \mathbf{v}; \psi) \parallel \mathrm{p}(\mathbf{z})\right]\right] + \mathrm{D}_{\mathrm{KL}}\left[\mathrm{q}(\mathbf{v} \mid \mathbf{u}; \psi) \parallel \mathrm{p}(\mathbf{v})\right].$$

Unsupervised Disentanglement of Anatomy and Geometry



Experiments
Multi-Modality & Inter-Subject Groupwise Registration

Datasets

- MS-CMRSeg: T1w, T2w, and LGE cardiac 2D MRI, intra-subject.
- BraTS-2021: T1w, T1ce, T2w, and FLAIR 3D brain MRI, intra-subject.
- Learn2Reg Abdomen MR-CT: MR and CT 3D abdomen images, intra/inter-subject.
- OASIS: T1w 3D brain MRI, inter-subject.

Experiments

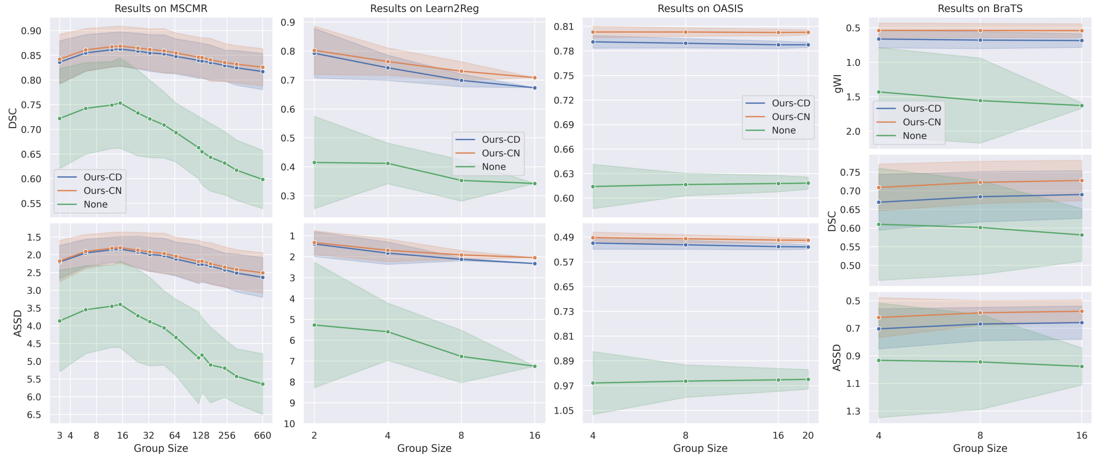
Multi-Modality & Inter-Subject Groupwise Registration

Method	MS-CMRSeg				BraTS-2021				
	DSC ↑	ASSD ↓	#Params	p-value	gWI ↓	DSC ↑	ASSD ↓	#Params	p-value
None	$.722 \pm .104$	3.86 ± 1.43	N/A	$< 10^{-10}$	1.430 ± 0.644	$.610 \pm .150$	0.93 ± 0.42	N/A	$< 10^{-10}$
APE [4]	$.746 \pm .101$	3.48 ± 1.36	0.154	$< 10^{-10}$	0.629 ± 0.141	$.707 \pm .069$	$\mathbf{0.62 \pm 0.12}$	7.37	1.3×10^{-1}
CTE [5]	$.766 \pm .100$	3.15 ± 1.32	0.154	$< 10^{-10}$	1.223 ± 0.255	$.500 \pm .102$	1.27 ± 0.45	7.37	$< 10^{-10}$
$\mathcal{X}$-CoReg [6]	$.757 \pm .107$	3.31 ± 1.40	0.154	$< 10^{-10}$	0.728 ± 0.196	$.698 \pm .086$	0.65 ± 0.17	7.37	$< 10^{-10}$
APE-Att	$.799 \pm .061$	2.78 ± 0.72	8.04	$< 10^{-10}$	0.757 ± 0.153	$.690 \pm .078$	0.67 ± 0.14	22.95	$< 10^{-10}$
CTE-Att	$.820 \pm .066$	2.45 ± 0.74	8.04	$< 10^{-10}$	0.916 ± 0.210	$.661 \pm .094$	0.75 ± 0.20	22.95	$< 10^{-10}$
Ours-PN	$.803 \pm .062$	2.68 ± 0.70	5.06 (5.22)	$< 10^{-10}$	0.608 ± 0.115	$.670 \pm .071$	0.69 ± 0.13	14.91 (15.10)	$< 10^{-10}$
Ours-CN	$\mathbf{.842 \pm .051}$	$\mathbf{2.17 \pm 0.58}$	5.06 (5.22)	$< 10^{-10}$	$\mathbf{0.538 \pm 0.108}$	$\mathbf{.709 \pm .063}$	0.62 ± 0.15	14.91 (15.10)	$< 10^{-10}$
Ours-PD	$.799 \pm .060$	2.74 ± 0.65	1.91 (2.85)	$< 10^{-10}$	0.720 ± 0.151	$.670 \pm .072$	0.70 ± 0.14	5.44 (15.06)	$< 10^{-10}$
Ours-CD	$.836 \pm .043$	2.21 ± 0.47	1.91 (2.85)	N/A	0.663 ± 0.129	$.669 \pm .074$	0.70 ± 0.15	5.44 (15.06)	N/A

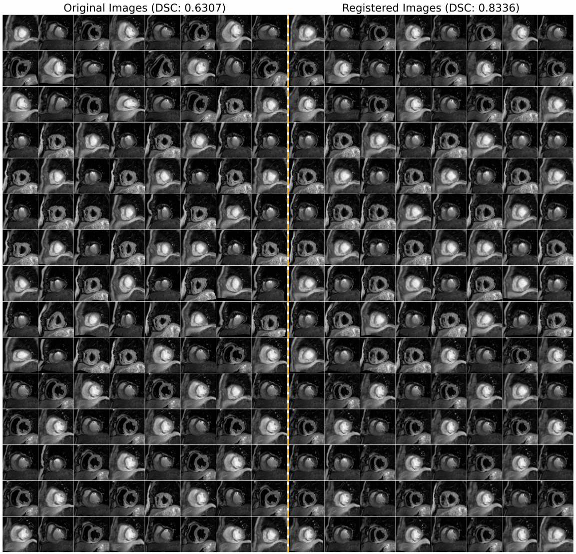
Method	Learn2Reg				OASIS				
	DSC ↑	ASSD ↓	#Params	p-value	$\lvert \det J_\phi \leq 0 \rvert$ (%) ↓	DSC ↑	ASSD ↓	#Params	p-value
None	$.415 \pm .160$	5.27 ± 3.01	N/A	2.4×10^{-4}	N/A	$.614 \pm .027$	0.96 ± 0.10	N/A	$< 10^{-10}$
APE [4]	$.554 \pm .384$	6.14 ± 7.66	10.52	9.7×10^{-2}	$.5297 \pm .0814$	$.777 \pm .030$	0.59 ± 0.09	8.85	7.9×10^{-8}
CTE [5]	$.514 \pm .373$	6.85 ± 7.89	10.52	5.1×10^{-2}	$.2291 \pm .0447$	$.746 \pm .031$	0.62 ± 0.09	8.85	$< 10^{-10}$
$\mathcal{X}$-CoReg [6]	$.664 \pm .361$	6.86 ± 14.1	10.52	3.0×10^{-1}	$.1330 \pm .0404$	$.777 \pm .017$	0.54 ± 0.05	8.85	$< 10^{-10}$
APE-Att	$.687 \pm .092$	2.09 ± 0.57	22.95	2.9×10^{-3}	$.0479 \pm .0130$	$.777 \pm .018$	0.57 ± 0.05	22.95	$< 10^{-10}$
CTE-Att	$.679 \pm .069$	2.17 ± 0.56	22.95	1.4×10^{-2}	$.0552 \pm .0154$	$.773 \pm .039$	0.59 ± 0.13	22.95	4.2×10^{-6}
Ours-CN	$\mathbf{.803 \pm .084}$	$\mathbf{1.32 \pm 0.57}$	14.90 (15.10)	9.3×10^{-2}	$.1746 \pm .0305$	$\mathbf{.803 \pm .007}$	$\mathbf{0.49 \pm 0.02}$	13.16 (13.26)	$< 10^{-10}$
Ours-CD	$.793 \pm .086$	1.39 ± 0.60	5.43 (15.06)	N/A	$\mathbf{.0066 \pm .0029}$	$.791 \pm .008$	0.51 ± 0.02	3.69 (13.22)	N/A

Experiments

Large-scale and Variable-size Image Groups

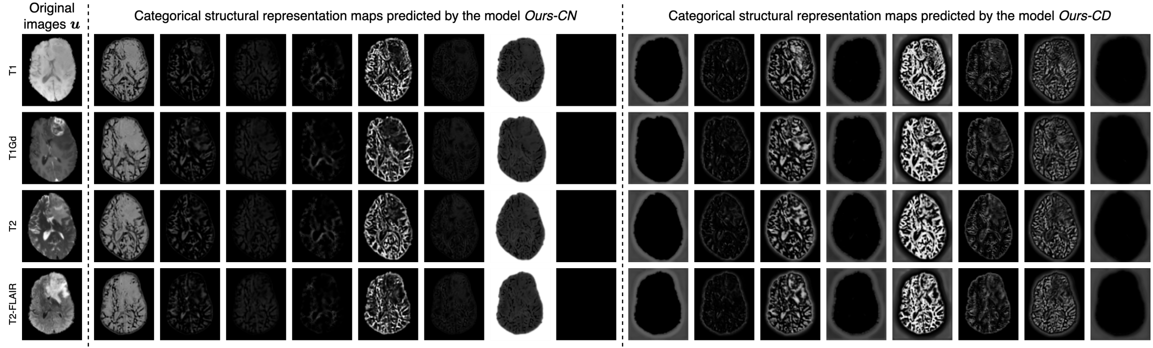


Evaluation metrics (mean values with one standard deviation bands) of registration results on image groups with different sizes.



An example of co-registration by our model (Ours-CD) for a group of 120 images from the MS-CMRSeg dataset.

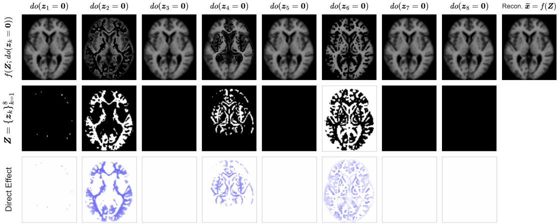
Model Interpretability
Structural Representation Maps



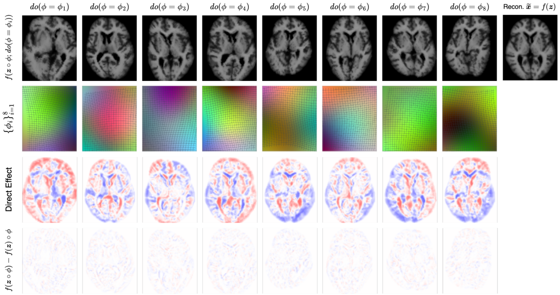
Categorical structural representation maps extracted from the proposed models. One can see that complementary brain structures are revealed from these representations.

Model Interpretability

Latent Symmetry Structures



Counterfactual reconstruction by ontological transformations.



Counterfactual reconstruction by diffeomorphic transformations.

Publications

- Luo, Xinzhe, and Xiahai Zhuang. “$\mathcal{X}$-Metric: An N-Dimensional Information-Theoretic Framework for Groupwise Registration and Deep Combined Computing.” IEEE TPAMI (2023) Link.
- Xinzhe Luo*, Xin Wang*, Linda Shapiro, Chun Yuan, Jianfeng Feng, Xiahai Zhuang. “Bayesian Intrinsic Groupwise Image Registration: Unsupervised Disentanglement of Anatomy and Geometry.” arXiv:2401.02141 (In Submission). Link.
- Xin Wang*, Xinzhe Luo*, and Xiahai Zhuang. “Bayesian intrinsic groupwise registration via explicit hierarchical disentanglement.” IPMI 2023 (Oral Presentation, Honorable Mention for the Francois Erbsmann Prize). Link.

* denotes equal contribution.

IMPERIAL

Thank you!
Questions?