This CVPR paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the accepted version;
the final published version of the proceedings is available on IEEE Xplore.

Learning Diffeomorphism for Medical Image Registration with Time-Embedded Architectures Using Semigroup Regularization

Mohammadjavad Matinkia
University of Alberta
Edmonton, Canada
matinkia@ualberta.ca

Nilanjan Ray
University of Alberta
Edmonton, Canada
nray1@ualberta.ca

Abstract

*Diffeomorphic image registration (DIR) seeks topology-preserving transformations and is fundamental in medical imaging. Existing DIR methods rely on integration schemes (e.g., scaling-and-squaring) and multiple regularizers to enforce invertibility. We introduce **SGDIR**, a continuous-time registration framework, parameterized by known time-embedded backbones, that models diffeomorphisms using only a single semigroup-based regularization, eliminating explicit integration and auxiliary constraints. We mathematically prove that this formulation directly learns the flow of an underlying ODE, inherently enforcing inverse and cycle consistencies. We evaluate on eight 2D and 3D MR and CT datasets. Under strict semigroup enforcement, our model achieves near-perfect diffeomorphism and significantly outperforms existing diffeomorphic methods, while remaining competitive with leading non-diffeomorphic deformable models. When the regularization is relaxed, the same architecture functions as a deformable method and substantially surpasses state-of-the-art non-diffeomorphic approaches in registration accuracy. These results demonstrate that continuous-time deformation modeling, guided solely by our semigroup-based regularization, yields a unified framework capable of both rigorously diffeomorphic mapping and state-of-the-art deformable registration*[1].

1. Introduction

Diffeomorphic image registration (DIR) seeks smooth, invertible transformations between images, preserving anatomical topology - an essential constraint in medical imaging where folding or tearing leads to invalid mappings [24, 46, 51].

DIR methods have evolved across three paradigms: traditional, instance-optimization, and learning-based. Traditional approaches such as SyN [3], NiftyReg [33], B-Splines [39], and LDDMM [6] optimize variational objectives under explicit smoothness constraints, achieving high accuracy but at prohibitive computational cost. Instance-optimization methods replace analytic models with neural networks [21, 43, 47, 54, 55, 59], improving efficiency while maintaining case-specific optimization. Fully learning-based models [4, 5, 14, 15, 34, 36, 42, 57] further amortize registration across datasets, enabling rapid inference.

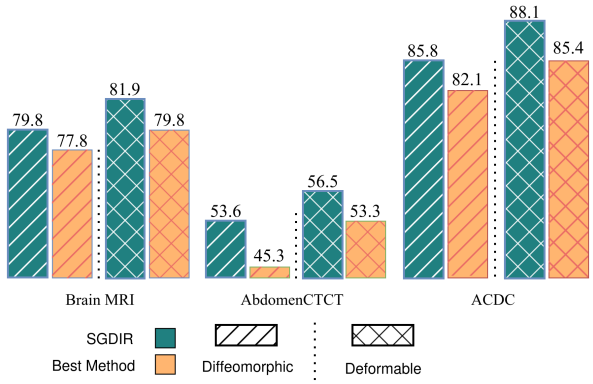


Figure 1. A performance summary of SGDIR in diffeomorphic and non-diffeomorphic settings comparing with best performing (non)diffeomorphic methods in the experiments. SGDIR shows on-par performance with top deformable models in diffeomorphic setting and outperforms them in non-diffeomorphic setting.

Most modern DIR frameworks parameterize velocity fields and reconstruct deformations via scaling-and-squaring integration [1, 22], coupled with multiple handcrafted regularizers (e.g., Jacobian penalties, smoothness constraints). Time-embedded architectures [23, 37, 38, 45] and diffusion-inspired models such as DiffuseMorph and DiffuseReg [28, 58] introduce continuous-time deformation, but still rely on auxiliary constraints rather than deriving diffeomorphism from first principles. Furthermore, inverse and cycle consistency - critical for medical registration - are often enforced explicitly rather than being inher-

[1] https://mattkia.github.io/SGDIR/

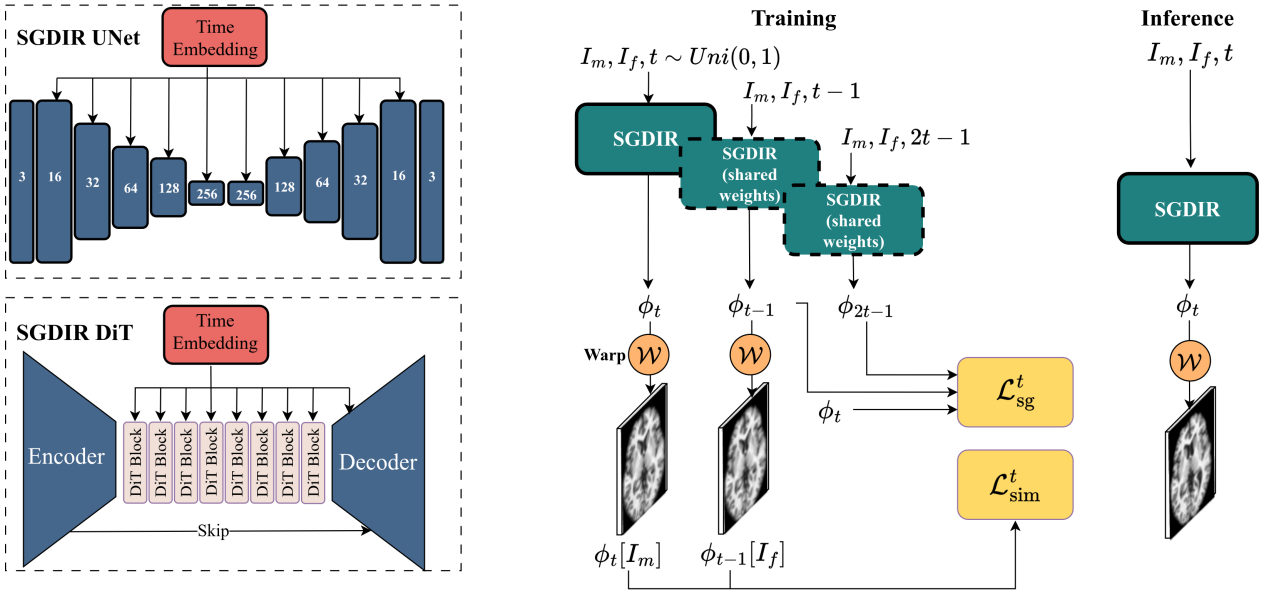


Figure 2. Architectures of SGDIR (left), Training scheme (middle) and inference scheme (right) of SGDIR. DiT blocks are implemented according to [37]. More details on the architecture can be found in the supplementary materials (Sec. 9).

ent to the model [12, 13, 16, 27]. Unlike prior ODE-based registration, we do not assume a velocity ODE, rather we derive it emergently from semigroup consistency.

We propose **SGDIR (Semigroup Diffeomorphic Image Registration)**, a continuous-time DIR framework that enforces diffeomorphism using only a single semigroup-based regularization, without any integration scheme or auxiliary constraints. Rather than imposing the full semigroup across all time pairs, SGDIR enforces a *partial* semigroup constraint (see Sec. 4), and we prove that this weaker condition alone is sufficient to guarantee that the model learns the flow of an ODE. This inductive bias inherently ensures invertibility, cycle consistency, and topology preservation. Unlike existing methods, SGDIR produces temporally continuous deformations at arbitrary time steps without scaling-and-squaring or smoothness penalties.

To instantiate SGDIR, we adopt time-embedded neural architectures commonly used in diffusion models, such as temporal UNets [23, 45] and Diffusion Transformer backbones [37], allowing continuous evaluation of deformation at arbitrary time points. However, our goal is not to introduce a new architecture; rather, we show that even standard backbones become inherently diffeomorphic when trained under our semigroup constraint. Thus, SGDIR is architecture-agnostic and decoupled from integration schemes or handcrafted regularizers.

Our contributions can be summarized as the followings:

- We introduce a continuous-time registration framework that enforces diffeomorphism using only the semigroup property, removing scaling-and-squaring and all auxiliary regularizers.
- We mathematically prove that our partial semigroup regularization compels the model to learn an ODE flow, inherently yielding inverse and cycle consistency and near-perfect diffeomorphism.
- Through a single framework, strict enforcement yields remarkable diffeomorphic stability (near-zero folding), while relaxed enforcement becomes a deformable model that surpasses state-of-the-art non-diffeomorphic methods.
- Experiments across eight 2D and 3D MR/CT datasets demonstrate that SGDIR outperforms existing diffeomorphic methods and significantly improves topology preservation over deformable methods, without sacrificing accuracy.

2. Related Works

Traditional deformable registration methods estimate spatial transformations by optimizing energy functionals with regularity constraints. Classical approaches such as SyN [3] and NiftyReg [33] optimize similarity metrics including cross-correlation, Normalized Mutual Information (NMI), and Sum of Squared Differences (SSD). Optical flow-based methods like Demons [48, 52, 53] and its variants further refine displacement fields through iterative updates. Flow-based formulations, notably LDDMM [6], model registration as geodesic shooting on diffeomorphism manifold, ensuring topology preservation under large deformations. DARTEL [2] similarly employs Lie algebra parameterizations and scaling-and-squaring for diffeomorphic integra-

tion. More recently, neural optimization methods using implicit representations, such as IDIR [54], DNVF [21], and NePhi [50], learn deformation or velocity fields directly in continuous coordinate spaces with Jacobian and hyperelastic regularization. NODEO [55] utilizes neural ODEs, integrating learned velocity dynamics via continuous-time solvers.

Learning-based registration shifts from pairwise optimization to data-driven prediction across large datasets, achieving orders-of-magnitude faster inference. VoxelMorph [5] pioneered CNN-based registration using UNets with localized NCC and smoothness losses. SYMNet [34] introduced symmetric velocity learning with scaling-and-squaring for diffeomorphic guarantees, while LapIRN, cLapIRN, and PULPo [35, 44] leveraged multi-resolution refinement to enhance generalization. Transformer-based models such as TransMorph, TransMatch, and H-ViT [9, 10, 20] replace convolutional backbones with attention mechanisms for improved long-range correspondence. CorrMLP [31] cleverly uses MLP decoders along with correlation maps for enhanced and lightweight registration. Diffusion-inspired frameworks like DiffuseMorph and DiffuseReg [28, 58] adopt denoising diffusion processes to learn temporally continuous deformation fields, though they require iterative sampling and fixed time grids, leading to high computational cost. In contrast, our approach employs time-embedded architectures (UNets and diffusion-style transformers) to learn continuous-time deformation fields that are instantly queryable at arbitrary timesteps, without iterative sampling.

3. Background

Deformable registration seeks a spatial transformation $\phi : \Omega \to \mathbb{R}^3$ that warps a moving image $I_m : \Omega \to \mathbb{R}$ into alignment with a fixed image $I_f : \Omega \to \mathbb{R}$. In diffeomorphic image registration (DIR), ϕ must be a smooth, orientation-preserving bijection, typically enforced via a positive Jacobian determinant $|J_\phi| > 0$ to avoid folding or tearing. Modern DIR frequently interprets ϕ as the endpoint of a continuous-time flow $\{\phi_t\}_{t\in[0,1]}$, with $\phi_0 = \mathrm{Id}$ and ϕ_1 used for warping I_m to I_f.

This flow is commonly modeled as the solution to an autonomous ordinary differential equation (ODE) [2, 6, 17, 32, 34]:

$$\begin{cases} \dfrac{d\phi_t}{dt} = \mathbf{v}(\phi_t), \\ \phi_0(x) = x, \end{cases} \tag{1}$$

where $\mathbf{v} : \mathbb{R}^3 \to \mathbb{R}^3$ is a stationary velocity field. The ODE formulation guarantees diffeomorphic maps under mild smoothness of $\mathbf{v}$. In practice, DIR methods estimate $\mathbf{v}$ and recover ϕ_1 using the classical *scaling-and-squaring* scheme. Starting from a small-step approximation

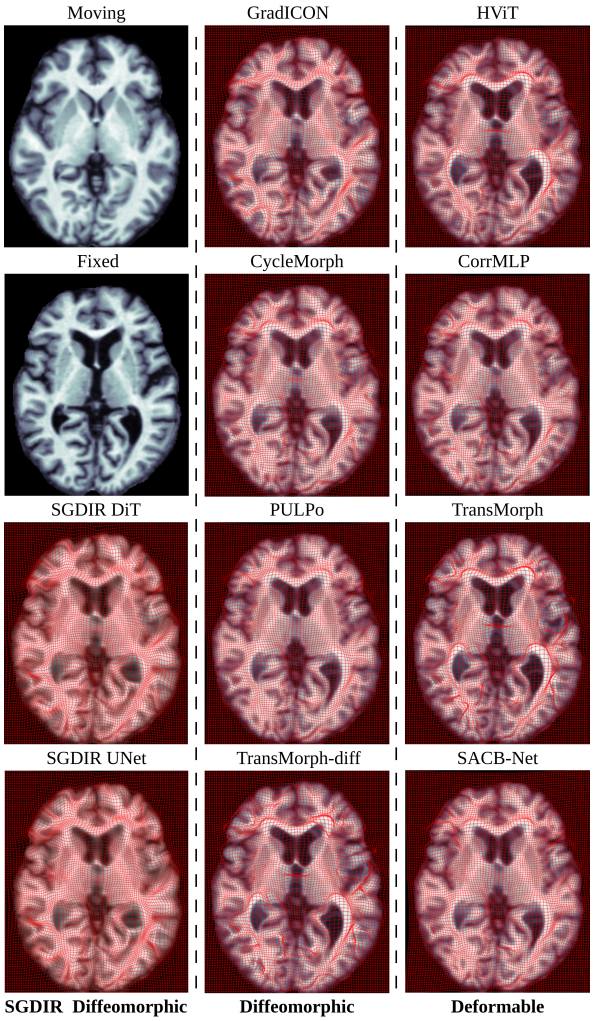


Figure 3. Visual comparison of SGDIR performance on the OASIS dataset with top diffeomorphic and non-diffeomorphic methods. The deformation grids are overlaid on the warped images. SGDIR exhibits the smoothest grids with high Dice accuracy.

$$\phi_{\frac{1}{2^N}}(x) = x + \frac{\mathbf{v}(x)}{2^N}, \tag{2}$$

successive composition yields

$$\phi_{\frac{1}{2^{N-1}}} = \phi_{\frac{1}{2^N}} \circ \phi_{\frac{1}{2^N}}, \quad \ldots, \quad \phi_1 = \phi_{\frac{1}{2}} \circ \phi_{\frac{1}{2}}, \tag{3}$$

which is embedded into the standard registration loss

$$\mathcal{L} = \mathbb{E}_{(I_f, I_m)\sim\mathcal{D}} \left[\mathcal{L}_{\text{sim}}(I_f, \phi_1[I_m]) + \mathcal{L}_{\text{reg}}(\phi)\right], \tag{4}$$

where $\mathcal{L}_{\text{sim}}$ is a measure of similarity and $\mathcal{L}_{\text{reg}}$ denotes the regularization terms. Also, $\phi[I_m]$ denotes the warping of I_m by ϕ. A key observation is that Eq. (1) is *equivalent* to

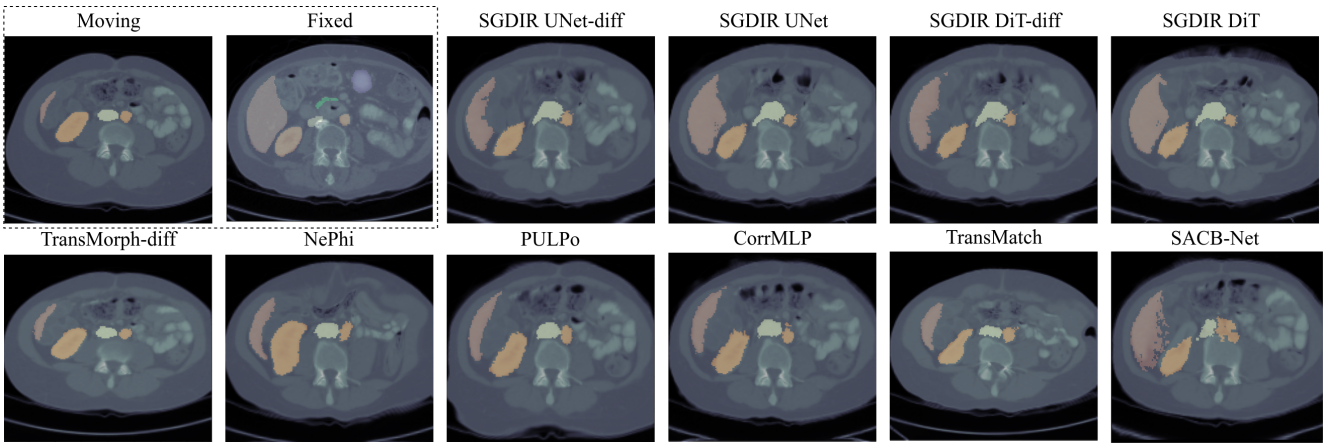


Figure 4. Visual comparison of SGDIR performance on the AbdomenCTCT dataset with top diffeomorphic and non-diffeomorphic methods. The *-diff* suffix denotes the SGDIR trained with $\lambda = 10^5$, and unsuffixed SGDIRs are trained with $\lambda = 10^4$.

enforcing the *semigroup property* on the deformation family $\{\phi_t\}$ [8]:

$$\phi_0 = \mathrm{Id}, \qquad \phi_t \circ \phi_s = \phi_{t+s}, \ \forall t, s. \tag{5}$$

Semigroup closure ensures temporal consistency: deforming to time $t + s$ is identical to composing deformations at times t and s. Moreover, if each ϕ_t is differentiable, Eq. (5) implies the existence of inverses via

$$\phi_t^{-1} = \phi_{-t}, \tag{6}$$

which guarantees bijectivity and diffeomorphic structure. In the next section, we show how enforcing only a *partial* semigroup constraint allows us to bypass scaling-and-squaring, eliminate auxiliary regularization, and learn ϕ_t directly in continuous time.

4. Proposed Method

Let $I_f, I_m : \Omega \longrightarrow \mathbb{R}$ be the fixed and moving images, respectively, sampled from the image distribution $\mathcal{D}$. The domain $\Omega \subset \mathbb{R}^3$ is a $D \times H \times W$ cubical region. We seek a time-continuous deformation field $\phi_t : \Omega \longrightarrow \mathbb{R}^3$, $t \in [-1, 1]$, $\phi_0(x) = x$ that transports I_m toward I_f, and vice versa. We assume that ϕ_t is the flow of an autonomous ODE as in Eq. (1) induced by an unknown stationary velocity field $\mathbf{v}$, and therefore satisfies the semigroup property of Eq. (5). By continuity of trajectories, warping the moving image forward to time t and the fixed image backward to time $1 - t$ must yield corresponding points. Using the inverse identity of Eq. (6) yields

$$\phi_{1-t}^{-1}[I_f] = \phi_{t-1}[I_f] = \phi_t[I_m]. \tag{7}$$

Adopting normalized cross-correlation (NCC) as similarity, Eq. (7) gives the following time-continuous similarity loss.

$$\mathcal{L}_{\text{sim}}^t = -NCC(\phi_{t-1}[I_f], \phi_t[I_m]), \qquad \forall t \in [0, 1]. \tag{8}$$

In practice, we parameterize $\phi_t(x)$ as

$$\phi_t(x; \theta) = \phi(x, t; \theta) = x + t\mathbf{F}(x, t; I_f, I_m, \theta), \tag{9}$$

for all $t \in [-1, 1]$, where $\mathbf{F}(.; \theta)$ is a time-embedded neural architecture with parameters θ, conditioned on both images. In our experiments, $\mathbf{F}(.; \theta)$ is instantiated using different architectures including time-embedded UNet [23, 45] and Diffusion Transformers [37]. Eq. (9) guarantees $\phi_0(x) = x$, satisfying the identity initial condition of the Eq. (1).

To ensure that ϕ_t forms a valid flow map, it must satisfy the semigroup law of Eq. (5). Enforcing the full law across arbitrary time pairs (t, s) is intractable in practice. We therefore introduce a tractable *partial* semigroup regularization,

$$\mathcal{L}_{\text{sg}}^t = \|\phi_{2t-1} - \phi_t \circ \phi_{t-1}\|_2 + \|\phi_{2t-1} - \phi_{t-1} \circ \phi_t\|_2, \tag{10}$$

where the second term enforces symmetry, and compositions are implemented via bilinear interpolation of deformation grids [21, 34–36]. We mathematically show that Eq. (10) is sufficient to make ϕ_t a flow map at all $t \in [-1, 1]$.

A minimizer of Eq. (10) satisfies

$$\phi_{2t-1} = \phi_t \circ \phi_{t-1}, \ \phi_{2t-1} = \phi_{t-1} \circ \phi_t. \tag{11}$$

We propose the following theorem and postpone the proof to appendix (Sec. 7):

		Atlas-Based				Inter-Subject			
Type	**Method**	**Dice↑**	$\lvert J\rvert_{<0}\%$↓	**HD95↓**	**SSIM↑**	**Dice↑**	$\lvert J\rvert_{<0}\%$↓	**HD95↓**	**SSIM↑**
Diffeomorphic	CycleMorph	82.07 ± 1.20	0.0084 ± 0.0123	2.79 ± 0.51	87.57 ± 1.49	81.58 ± 2.03	0.0173 ± 0.0086	2.26 ± 0.79	88.10 ± 1.38
	GradICON	84.53 ± 1.52	0.0022 ± 0.0019	2.23 ± 0.55	85.90 ± 1.65	83.89 ± 1.82	0.0039 ± 0.0012	2.09 ± 0.36	90.17 ± 1.19
	TransMorph-diff	84.63 ± 1.12	0.0091 ± 0.0115	2.25 ± 0.53	89.91 ± 1.13	83.51 ± 1.52	0.0066 ± 0.0073	2.35 ± 0.76	91.25 ± 2.07
	R2Net	80.78 ± 1.47	0.0036 ± 0.0028	2.63 ± 0.74	86.14 ± 1.52	80.11 ± 1.53	0.0093 ± 0.0054	2.15 ± 0.49	87.44 ± 1.15
	NODEO	81.87 ± 3.87	0.0016 ± 0.0020	2.38 ± 0.99	87.05 ± 2.69	81.97 ± 1.39	0.0024 ± 0.0010	2.18 ± 0.71	87.68 ± 1.93
	SGDIR DiT ($\lambda = 10^5$)	**86.53 ± 0.82**	**0.0 ± 0.0**	**1.90 ± 0.52**	**91.45 ± 1.01**	**85.90 ± 1.94**	0.0003 ± 0.0005	**1.87 ± 0.59**	**92.52 ± 0.97**
	SGDIR UNet ($\lambda = 10^5$)	86.16 ± 0.89	**0.0 ± 0.0**	1.96 ± 0.51	90.71 ± 1.01	85.36 ± 1.67	**0.0001 ± 0.0002**	1.93 ± 0.47	92.12 ± 1.47
Deformable	TransMorph	85.26 ± 1.05	2.0155 ± 0.4260	2.39 ± 0.49	91.79 ± 1.90	84.11 ± 1.30	1.0665 ± 0.5631	2.26 ± 0.68	91.14 ± 1.49
	TransMatch	81.54 ± 1.45	0.0237 ± 0.0154	2.37 ± 0.57	88.22 ± 1.37	83.48 ± 1.80	0.1493 ± 0.0151	2.23 ± 0.66	90.26 ± 2.11
	HViT	85.38 ± 0.93	0.3566 ± 0.0832	2.13 ± 0.54	92.02 ± 1.07	85.07 ± 1.05	0.4812 ± 0.0114	1.92 ± 0.46	92.13 ± 1.39
	CorrMLP	83.07 ± 0.91	0.0125 ± 0.0059	2.13 ± 0.55	88.32 ± 1.88	84.66 ± 1.24	0.4640 ± 0.1776	2.23 ± 0.62	90.29 ± 1.84
	SACB-Net	81.60 ± 0.61	**0.0020 ± 0.0013**	2.19 ± 0.48	86.82 ± 0.94	83.17 ± 1.18	**0.0184 ± 0.0093**	2.33 ± 0.48	88.66 ± 2.13
	SGDIR DiT ($\lambda = 10^4$)	**88.09 ± 0.91**	0.4332 ± 0.1121	**1.73 ± 0.41**	**93.80 ± 1.15**	**87.82 ± 0.77**	0.3982 ± 0.1213	1.83 ± 0.65	**95.32 ± 1.08**
	SGDIR UNet ($\lambda = 10^4$)	87.97 ± 0.85	0.5991 ± 0.1043	1.96 ± 0.48	93.44 ± 0.80	87.74 ± 1.13	0.5018 ± 0.0243	**1.63 ± 0.33**	94.60 ± 1.25

Table 1. Quantitative registration results on the OASIS dataset in atlas-based and inter-subject settings. The best result is shown in **bold** and the best competitor is presented in blue.

Theorem 1. *The deformation $\phi_t(x)$ satisfying the composition rule of Eq. (11) and $\phi_0(x) = x$ is an exponential map, or equivalently, a one-parameter diffeomorphism solving the autonomous ODE of Eq. (1); see [7].*

Following Eq. (4), the overall training objective is defined as

$$\mathcal{L} = \mathbb{E}_{(I_f, I_m)\sim\mathcal{D},\, t\sim\mathrm{Uni}(0,1)} \left[\mathcal{L}^t_{\mathrm{sim}} + \lambda \mathcal{L}^t_{\mathrm{sg}}\right] \tag{12}$$

where λ controls the strength of diffeomorphic regularization. Larger λ enforces stricter flow consistency, whereas smaller values allow deformable flexibility. During training, we uniformly sample image pairs and time values to minimize (12). The trained network can warp either image toward the other at arbitrary time t, in contrast to prior methods that predict deformation only at $t = 1$.

Finally, note that earlier approaches [12, 13, 27] impose consistency only at the endpoint by explicitly predicting both forward and backward fields at ϕ_1 and ϕ_1^{-1}. In contrast, our semigroup formulation yields bidirectional consistency for all $t \in [-1, 1]$ without explicit inverse modeling.

5. Experiments

5.1. Datasets and Metrics

We evaluate the proposed SGDIR framework on eight publicly available 2D and 3D medical imaging datasets commonly used in registration benchmarks, including OASIS [30], CANDI [26], LPBA40 [41], Mindboggle101 [29], IXI [2], ACDC [40], and the Learn2Reg LungCT and AbdomenCTCT datasets [3]. These datasets comprise T1-weighted brain MRIs, cardiac MRIs, and thoracic and abdominal CT scans, covering diverse anatomies and modalities. Quantitative evaluation is performed using the Dice score (for datasets with available segmentation maps), target registration error (TRE; for datasets with annotated landmarks), 95th percentile Hausdorff Distance (HD95), structural similarity index measure (SSIM), average symmetric surface distance (ASSD), and the percentage of voxels with negative Jacobian determinants. Dataset introductions and preprocessings are detailed in the supplementary materials (Sec. 8).

5.2. Experimental Settings

For all datasets, we adopt standard training, validation, and test splits consistent with prior works. Detailed dataset splitting, subject counts and pair generation protocols for all datasets are provided in the supplementary materials (Sec. 8). To ensure fair and transparent evaluation within page constraints, we report results for the top five performing baselines in the main text, while full comparison tables including all methods are provided in the supplementary materials (Sec. 10). On brain MRI datasets, we present the complete set of evaluation metrics for OASIS, and report the main metrics (Dice and $\lvert J\rvert_{<0}\%$) for the remaining datasets in the main text, with full results available in the supplementary materials (Sec. 10).

We evaluate four variants of SGDIR: two diffeomorphic models ($\lambda = 10^5$) using time-embedded UNet and Diffusion Transformer (DiT) backbones, and two deformable counterparts ($\lambda = 10^4$) with the same architectures (Fig. 2). We compare SGDIR against a comprehensive set of registration paradigms, including classical diffeomorphic optimization (SyN [3]), learning-based diffeomorphic models (SYMNet [34], LapIRN [35], NePhi [50], PULPo [44], Transmorph-diff [9], CycleMorph [27], GradICON [49],

[2] https://brain-development.org/ixi-dataset/

[3] https://learn2reg.grand-challenge.org/Datasets/

Type	Method	IXI		Mindboggle101		LPBA40		CANDI	
		Dice↑	$\lvert J\rvert_{<0}\%\downarrow$	Dice↑	$\lvert J\rvert_{<0}\%\downarrow$	Dice↑	$\lvert J\rvert_{<0}\%\downarrow$	Dice↑	$\lvert J\rvert_{<0}\%\downarrow$
Diffeomorphic	CycleMorph	74.67 ± 1.33	0.0224 ± 0.0110	70.64 ± 1.48	0.1282 ± 0.0873	71.90 ± 1.41	0.0064 ± 0.0032	81.37 ± 1.25	0.0027 ± 0.0013
	GradICON	76.43 ± 1.28	0.0018 ± 0.0022	70.78 ± 1.52	0.0117 ± 0.0043	72.21 ± 1.26	0.0011 ± 0.0006	83.39 ± 1.09	0.0015 ± 0.0008
	NePhi	75.34 ± 1.78	0.0009 ± 0.0016	67.91 ± 1.91	0.0009 ± 0.0005	72.39 ± 1.33	**0.0 ± 0.0**	79.15 ± 1.46	0.0009 ± 0.0005
	TransMorph-diff	75.98 ± 2.01	0.0142 ± 0.0098	69.45 ± 1.66	0.0364 ± 0.0157	70.98 ± 1.52	0.0086 ± 0.0044	83.20 ± 1.14	0.0026 ± 0.0012
	PULPo	73.91 ± 2.11	0.0066 ± 0.0043	68.69 ± 1.74	0.0287 ± 0.0104	74.82 ± 1.11	0.0009 ± 0.0005	78.38 ± 1.57	0.0136 ± 0.0047
	NODEO	75.68 ± 1.82	0.0011 ± 0.0015	70.18 ± 1.58	0.0023 ± 0.0011	72.73 ± 1.25	0.0008 ± 0.0004	80.60 ± 1.26	0.0012 ± 0.0007
	SGDIR DiT ($\lambda = 10^5$)	**80.18 ± 1.09**	**0.0 ± 0.0**	**71.58 ± 1.12**	**0.0004 ± 0.0003**	**77.13 ± 1.10**	**0.0 ± 0.0**	**84.50 ± 1.08**	0.0003 ± 0.0002
	SGDIR UNet ($\lambda = 10^5$)	80.04 ± 1.02	0.0002 ± 0.0001	71.20 ± 1.45	0.0005 ± 0.0002	76.83 ± 1.15	**0.0 ± 0.0**	84.36 ± 1.11	**0.0001 ± 0.0001**
Deformable	TransMorph	77.72 ± 1.76	1.2574 ± 0.7574	71.46 ± 1.63	1.4883 ± 0.5328	74.42 ± 1.33	0.5275 ± 0.1835	84.48 ± 1.10	0.1059 ± 0.0388
	TransMatch	76.80 ± 1.85	0.0406 ± 0.0120	71.78 ± 1.49	0.2483 ± 0.1020	74.13 ± 1.31	0.0251 ± 0.0116	82.60 ± 1.18	0.0160 ± 0.0061
	HViT	80.67 ± 1.67	0.5933 ± 0.1028	71.80 ± 1.72	1.4051 ± 0.6153	76.57 ± 1.18	0.4428 ± 0.1655	84.67 ± 1.07	0.1368 ± 0.0435
	CorrMLP	77.54 ± 1.68	0.3675 ± 0.2168	71.92 ± 1.68	**0.1895 ± 0.0782**	75.72 ± 1.27	0.0145 ± 0.0068	81.96 ± 1.25	0.0098 ± 0.0043
	SACB-Net	78.15 ± 1.87	**0.0201 ± 0.0099**	72.75 ± 1.57	0.3542 ± 0.1433	73.91 ± 1.35	**0.0099 ± 0.0041**	82.75 ± 1.22	**0.0047 ± 0.0021**
	DiffuseReg	72.13 ± 1.88	0.5279 ± 0.2834	71.88 ± 1.54	1.0823 ± 0.4882	72.98 ± 1.46	0.0723 ± 0.0289	79.31 ± 1.38	0.0322 ± 0.0125
	SGDIR DiT ($\lambda = 10^4$)	**83.88 ± 1.77**	0.6036 ± 0.0893	**73.59 ± 1.38**	0.6512 ± 0.0813	**78.93 ± 1.02**	0.1821 ± 0.0731	**85.51 ± 0.96**	0.0473 ± 0.0157
	SGDIR UNet ($\lambda = 10^4$)	83.73 ± 1.52	0.5803 ± 0.0988	73.18 ± 1.63	0.6143 ± 0.2441	77.09 ± 1.08	0.1730 ± 0.0691	85.28 ± 0.98	0.0331 ± 0.0132

Table 2. Quantitative registration results across MRI brain datasets. For clarity, only Dice and topology violation rates ($\lvert J\rvert_{<0}\%$) are shown. Full metrics including HD95, SSIM and ASSD are provided in the supplementary materials (Sec. 10). The best result is shown in **bold** and the best competitor is presented in blue.

R2Net [25]), and ODE-based approaches (NODEO [55]). We further include non-diffeomorphic deformable baselines (H-ViT [20], Transmorph [9], CorrMLP [31], DiffuseMorph [28], DiffuseReg [58], SACB-Net [11], AdaCS [56], TransMatch [10], VoxelMorph [5]).

SyN is evaluated via DIPY [19] package. All baselines are reproduced using official implementations and hyperparameters. All experiments are done using an NVIDIA RTX 3090 with 24GB VRAM. Full implementation details of SGDIR are deferred to supplementary materials (Sec. 9).

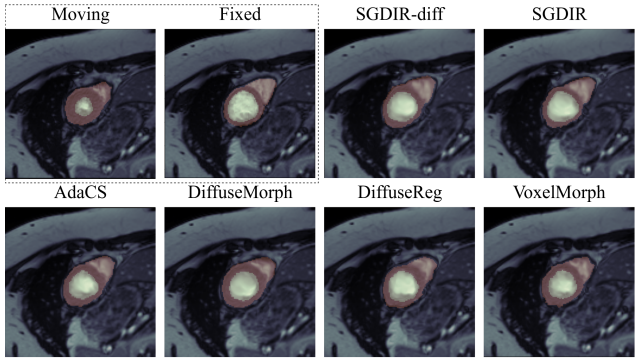


Figure 5. Visual comparisons of SGDIR performance on the ACDC dataset with overlaid segmentation masks.

5.3. Results

Tab. 1, and Tab. 2 report comprehensive quantitative comparisons on MRI brain scan datasets with available segmentation labels, and Tab. 3 presents the results on the AbdomenCTCT dataset. These tables provide the performance in terms of Dice, HD95, the percentage of negative Jacobian determinants ($\lvert J_{<0}\rvert\%$), SSIM, and ASSD. As mentioned before, Tab. 2 reports the main evaluation metrics, while the full set of metrics are reported in appendix (Sec. 10). Tab. 4 summarizes results on the LungCT dataset for which we report the performance in terms of TRE, $\lvert J_{<0}\rvert\%$, and SSIM. Finally, Tab. 5 provides results on the 2D ACDC dataset, where registration is performed between end-diastolic (ED) and end-systolic (ES) frames. For OASIS, we evaluate both atlas-based and inter-subject settings; all other datasets are evaluated in inter-subject mode. Metrics are expressed as mean ± standard deviation across the test set. Fig. 3, 4, and 5 illustrate examples of the registration results on the OASIS, AbdomenCTCT, and ACDC datasets, respectively. A comprehensive set of visual results can be found in the supplementary materials (Sec. 11).

Type	Method	Dice↑	$\lvert J\rvert_{<0}\%\downarrow$	HD95↓	ASSD↓
Diffeomorphic	CycleMorph	42.56 ± 2.65	0.0053 ± 0.0025	13.04 ± 1.28	4.77 ± 1.08
	GradICON	41.02 ± 2.81	0.0009 ± 0.0004	13.21 ± 1.33	4.86 ± 1.19
	NePhi	45.32 ± 2.24	0.0008 ± 0.0003	12.48 ± 1.19	3.90 ± 0.68
	TransMorph-diff	41.41 ± 2.72	0.0034 ± 0.0017	12.39 ± 1.25	4.71 ± 0.76
	PULPo	44.43 ± 2.13	0.0766 ± 0.0214	10.93 ± 1.04	3.51 ± 0.62
	SGDIR DiT ($\lambda = 10^5$)	52.23 ± 2.04	0.0001 ± 0.0001	10.27 ± 0.94	**2.89 ± 0.38**
	SGDIR UNet ($\lambda = 10^5$)	**53.64 ± 2.01**	**0.0 ± 0.0**	**10.07 ± 0.85**	2.97 ± 0.45
Deformable	TransMorph	46.66 ± 2.48	3.1310 ± 0.8113	12.61 ± 1.28	4.32 ± 1.12
	TransMatch	43.64 ± 2.35	2.6378 ± 0.6922	12.50 ± 1.22	3.87 ± 0.94
	HViT	47.66 ± 2.41	0.9601 ± 0.4175	10.11 ± 1.10	3.88 ± 0.83
	CorrMLP	50.28 ± 2.18	**0.1656 ± 0.0643**	11.47 ± 1.16	3.00 ± 0.53
	SACB-Net	53.38 ± 2.63	0.9348 ± 0.1135	13.09 ± 1.41	3.67 ± 0.49
	SGDIR DiT($\lambda = 10^4$)	54.15 ± 1.91	0.6244 ± 0.2184	10.19 ± 0.89	**2.45 ± 0.27**
	SGDIR UNet ($\lambda = 10^4$)	**56.57 ± 1.83**	0.2683 ± 0.0945	**9.19 ± 0.78**	2.45 ± 0.31

Table 3. Quantitative registration results on the AbdomenCTCT dataset. The best result is shown in **bold** and the best competing method is presented in blue.

Across Tab. 1–5, SGDIR under diffeomorphic regularization ($\lambda = 10^5$) consistently surpasses state-of-the-art diffeomorphic methods in Dice/TRE while matching the accuracy of competitive non-diffeomorphic approaches, and

Type	Method	TRE↓	$\lvert J\rvert_{<0}\%$↓	SSIM↑
Diffeomorphic	LapIRN	2.93 ± 0.18	0.0043 ± 0.0014	54.15 ± 3.05
	GradICON	2.64 ± 0.17	0.0009 ± 0.0004	54.06 ± 2.86
	TransMorph-diff	2.89 ± 0.18	0.0042 ± 0.0015	66.85 ± 2.95
	R2Net	2.99 ± 0.19	0.0211 ± 0.0068	68.06 ± 2.81
	NODEO	2.87 ± 0.18	0.0038 ± 0.0013	**70.80 ± 2.74**
	SGDIR DiT($\lambda = 10^5$)	2.77 ± 0.77	**0.0 ± 0.0**	66.06 ± 2.46
	SGDIR UNet ($\lambda = 10^5$)	**2.37 ± 0.15**	**0.0 ± 0.0**	66.55 ± 2.61
Deformable	TransMorph	2.85 ± 0.17	0.4767 ± 0.1546	68.82 ± 2.65
	TransMatch	3.02 ± 0.19	0.1343 ± 0.0482	67.47 ± 2.89
	HViT	3.22 ± 0.20	0.1791 ± 0.0554	46.72 ± 3.41
	CorrMLP	2.48 ± 0.16	0.2673 ± 0.0711	55.57 ± 3.08
	SACB-Net	3.01 ± 0.19	0.9824 ± 0.2761	70.06 ± 2.77
	SGDIR DiT($\lambda = 10^4$)	2.57 ± 0.65	**0.0543 ± 0.0187**	70.02 ± 2.66
	SGDIR UNet ($\lambda = 10^4$)	**2.23 ± 0.14**	0.0615 ± 0.0223	**70.48 ± 2.44**

Table 4. Quantitative registration results on the LungCT dataset.The best result is shown in **bold** and the best competing result is presented in blue.

Method	Dice↑	$\lvert J\rvert_{<0}\%$↓	HD95↓	SSIM↑	ASSD↓
AdaCS	85.46 ± 1.12	0.1493 ± 0.0185	5.56 ± 0.34	84.46 ± 1.47	2.12 ± 1.20
VoxelMorph	80.26 ± 1.28	0.0256 ± 0.0067	7.16 ± 0.42	85.12 ± 1.35	2.72 ± 1.73
DiffuseMorph	81.90 ± 1.35	0.0383 ± 0.0091	7.34 ± 0.47	79.81 ± 1.62	2.81 ± 1.48
DiffuseReg	83.31 ± 1.21	0.2345 ± 0.0228	6.17 ± 0.39	82.13 ± 1.54	2.35 ± 1.31
TransMorph	84.47 ± 1.18	0.0418 ± 0.0087	6.24 ± 0.36	80.08 ± 1.41	2.54 ± 1.22
TransMorph-diff	82.13 ± 1.42	0.0012 ± 0.0005	6.55 ± 0.43	79.56 ± 1.68	2.67 ± 1.36
SGDIR DiT($\lambda = 10^5$)	**85.79 ± 1.23**	**0.0 ± 0.0**	**5.15 ± 0.41**	**85.86 ± 1.79**	**1.94 ± 1.02**
SGDIR UNet ($\lambda = 10^5$)	85.49 ± 1.03	0.0003 ± 0.0001	5.22 ± 0.31	85.36 ± 1.29	2.01 ± 1.10
SGDIR DiT($\lambda = 10^4$)	**88.07 ± 1.13**	0.1277 ± 0.0312	**4.51 ± 0.49**	**87.49 ± 1.32**	**1.70 ± 0.82**
SGDIR UNet ($\lambda = 10^4$)	87.17 ± 0.97	0.1656 ± 0.0174	4.73 ± 0.28	86.91 ± 1.15	1.77 ± 1.03

Table 5. Quantitative registration results on the ACDC dataset. The best result is shown in **bold** and the best competitor is presented in blue.

significantly reduces topology violations ($\lvert J_{<0}\rvert\%$) on both MRI and CT modalities. When the semigroup constraint is relaxed ($\lambda = 10^4$), the deformable SGDIR variant achieves further gains, notably outperforming all baselines across modalities.

On MRI datasets, diffeomorphic SGDIR yields an average Dice improvement of $+2.5\%$ over leading diffeomorphic methods, and $+2.5\%$ compared to non-diffeomorphic deformable SOTA. On AbdomenCTCT, SGDIR achieves over $+5\%$ Dice improvement, and on LungCT it reduces TRE by approximately 10% relative to the best-performing models. A summary of performance can be seen in Fig. 1. In all experiments, the proportion of voxels with negative Jacobian determinants remains consistently low, indicating that enforcing semigroup consistency over continuous time effectively induces smooth and anatomically plausible deformations.

5.4. Ablation Study

We evaluate the influence of key design choices in SGDIR, focusing on: (i) the semigroup regularization weight λ, (ii) topology preservation over the time interval, (iii) the effect of discrete vs. continuous time sampling, and (iv) computational analysis of SGDIR. Since qualitative trends were consistent across datasets, we report all ablation results on the OASIS dataset for clarity.

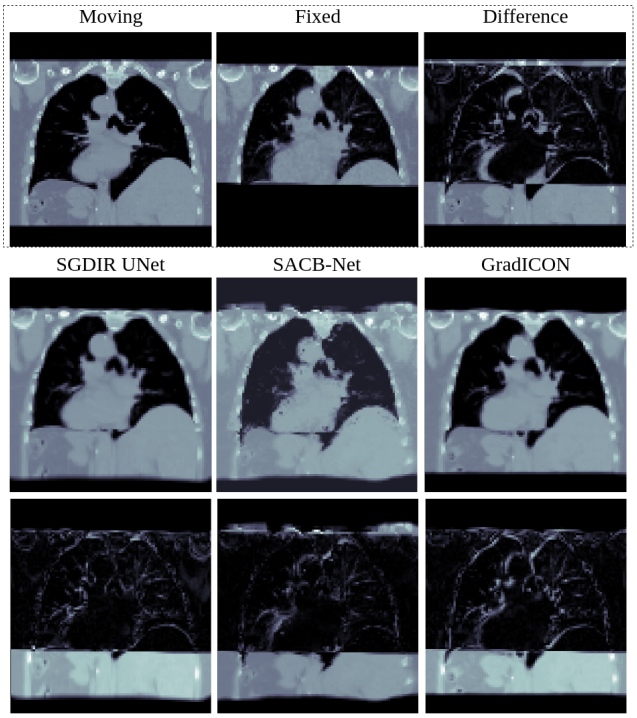


Figure 6. Visual comparison over LungCT dataset. The best performing SGDIR (left), deformable method (middle), and diffeomorphic method (right) are shown with absolute difference after registration.

	OASIS		LungCT	
λ	Dice ↑	$\lvert J\rvert_{<0}\%$ ↓	TRE ↓	$\lvert J\rvert_{<0}\%$ ↓
10^5	85.90	**0.0003**	2.37	**0.0**
10^4	**87.82**	0.3982	**2.23**	0.0615
10^3	84.66	3.1876	2.66	1.0183
10^2	81.01	6.7403	3.15	3.1098
10	80.23	7.0794	3.78	5.3951
0	79.80	7.8612	3.99	6.7632

Table 6. Effect of λ on the Dice, TRE and $\lvert J\rvert_{<0}\%$ over the OASIS and LungCT datasets.

5.4.1. Effect of Regularization Factor

The regularization term in Eq. (12) enforces adherence to the semigroup property, promoting topology-preserving deformations. Tab. 6 reports performance across different values of λ in terms of Dice, TRE and $\lvert J\rvert_{<0}\%$. A large λ enforces strict diffeomorphism, yielding zero folding but slightly reduced accuracy. Reducing λ allows greater flexibility and improves Dice/TRE (deformable SGDIR), but excessive relaxation degrades accuracy. This demonstrates that semigroup regularization not only prevents folding but also constrains the deformation trajectory, stabilizing optimization and guiding the model toward more anatomically valid solutions.

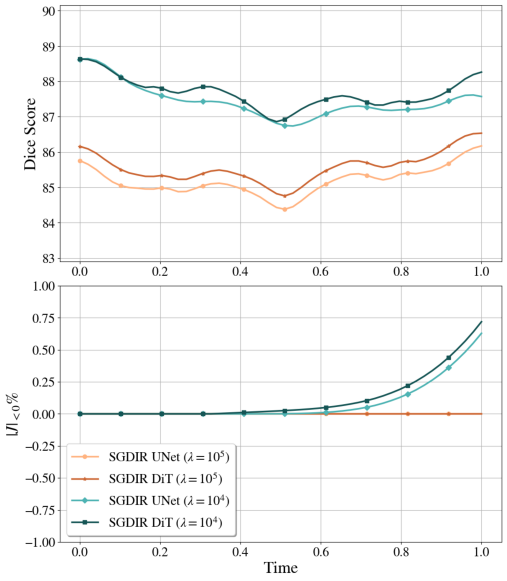


Figure 7. The Dice score (top) and topology preservation of forward deformation (bottom) of SGDIR variants throughout time.

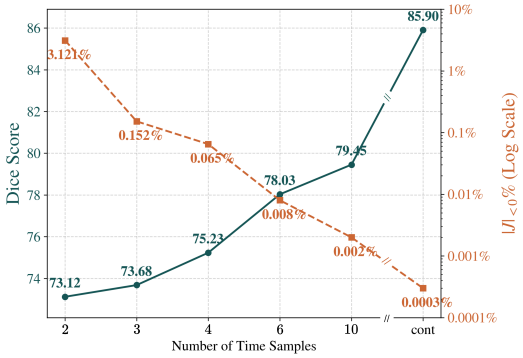


Figure 8. The effect of training with discrete vs. continuous (*cont*) time sampling on the inference for SGDIR DiT ($\lambda = 10^5$).

5.4.2. Diffeomorphism Through Time

Semigroup regularization enforces diffeomorphic consistency across the continuous time domain $t \in [0, 1]$. Fig. 7 visualizes average $|J|_{<0}\%$ (bottom) and Dice scores (top) across time. The Dice score at time t is computed by warping both fixed and moving images to the same time point. The bottom panel shows that SGDIR with diffeomorphic setting preserves topology perfectly throughout time, whereas deformable SGDIR only accumulates the folding near $t = 1$. The top panel indicates that SGDIR achieves a consistently high Dice score across all time steps, demonstrating smooth and invertible transformations without compromising alignment accuracy.

5.4.3. Effect of Discrete Time Sampling

SGDIR is trained using continuous time sampling. Here, we study the effect of replacing continuous sampling with a discrete set of equidistant time steps in $[0, 1]$. For example, using two time samples corresponds to training only at the endpoints $t = 0$ and $t = 1$. At evaluation, the model is queried at $t = 1$ to warp the moving image toward the fixed image. As shown in Fig. 8, increasing the number of discrete time samples consistently improves performance in both Dice score and $|J|_{<0}\%$, with the best results achieved under continuous sampling. In this figure, the term *cont* denotes the continuous sampling scheme used in our method.

5.4.4. Computational Analysis

We compare the computational efficiency of SGDIR against both a top-performing non-diffeomorphic transformer-based model (HViT) and a diffeomorphic integration-based model (TransMorph-diff). By eliminating expensing integration schemes, SGDIR achieves near $2\times$ faster inference speed and half the inference memory usage compared to TransMorph and HViT (Tab. 7).

	SGDIR UNet	SGDIR DiT	HViT	Transmorph-diff
# Params	**12.8M**	68.6M	21.2M	46.8M
Train Max Mem	22.6GB	16.12GB	23.2GB	**12.2GB**
Train Time/Iter	1.34s	1.16s	1.02s	**0.98s**
Test Max Mem	**2.82GB**	5.17GB	5.76GB	4.67GB
Test Time/Iter	0.25s	**0.22s**	0.51s	0.46s

Table 7. Computational analysis of SGDIR reveals the highest inference speed and lowest memory usage compared to top performing deformable and integration-based diffeomorphic methods.

6. Conclusion

We introduced **SGDIR** for diffeomorphic image registration using time-embedded architectures and partial semigroup regularization. We provided theoretical guarantees for the partial semigroup regularization to produce diffeomorphisms. SGDIR enables the generation of intermediate warps without numerical integration, while the semigroup constraint encourages both smoothness and invertibility of the resulting transformations. Our analysis revealed that strict regularization yields a state-of-the-art diffeomorphic model with minimal topology violations, whereas relaxing the constraint produces a highly flexible deformable variant that surpasses leading deformable methods. Unlike conventional approaches that rely on auxiliary smoothness or inverse consistency penalties, SGDIR achieves diffeomorphic behavior intrinsically through temporal consistency. Extensive experiments across eight 2D and 3D medical benchmarks demonstrated consistent improvements in accuracy, anatomical plausibility, and robustness. These results highlight SGDIR as a principled and versatile paradigm for diffeomorphic/deformable image registration.

Acknowledgments

The authors acknowledge funding support from NSERC Discovery Grants and the Department of Computing Science, University of Alberta. The project received compute support from Digital Research Alliance of Canada.

References

[1] Vincent Arsigny, Olivier Commowick, Xavier Pennec, and Nicholas 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. 1

[2] John Ashburner. A fast diffeomorphic image registration algorithm. *Neuroimage*, 38(1):95–113, 2007. 2, 3

[3] Brian B Avants, Charles L Epstein, Murray Grossman, and James 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. 1, 2, 5

[4] Guha Balakrishnan, Amy Zhao, Mert R Sabuncu, John Guttag, and Adrian V Dalca. An unsupervised learning model for deformable medical image registration. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 9252–9260, 2018. 1

[5] Guha Balakrishnan, Amy Zhao, Mert R Sabuncu, John Guttag, and Adrian V Dalca. Voxelmorph: a learning framework for deformable medical image registration. *IEEE transactions on medical imaging*, 38(8):1788–1800, 2019. 1, 3, 6

[6] M Faisal Beg, Michael I Miller, Alain Trouvé, and Laurent Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International journal of computer vision*, 61:139–157, 2005. 1, 2, 3

[7] S. Biagi and A. Bonfiglioli. *An Introduction to the Geometrical Analysis of Vector Fields: With Applications to Maximum Principles and Lie Groups*. World Scientific, 2018. 5, 2

[8] Stefano Biagi and Andrea Bonfiglioli. *An Introduction to the Geometrical Analysis of Vector Fields: with Applications to Maximum Principles and Lie Groups*. World Scientific, 2019. 4

[9] Junyu Chen, Eric C Frey, Yufan He, William P Segars, Ye Li, and Yong Du. Transmorph: Transformer for unsupervised medical image registration. *Medical image analysis*, 82:102615, 2022. 3, 5, 6

[10] Zeyuan Chen, Yuanjie Zheng, and James C Gee. Transmatch: a transformer-based multilevel dual-stream feature matching network for unsupervised deformable image registration. *IEEE transactions on medical imaging*, 43(1):15–27, 2023. 3, 6

[11] Xinxing Cheng, Tianyang Zhang, Wenqi Lu, Qingjie Meng, Alejandro F Frangi, and Jinming Duan. Sacb-net: Spatial-awareness convolutions for medical image registration. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 5227–5237, 2025. 6

[12] Gary E Christensen and Hans J Johnson. Consistent image registration. *IEEE transactions on medical imaging*, 20(7): 568–582, 2001. 2, 5

[13] Gary E Christensen, Joo Hyun Song, Wei Lu, Issam El Naqa, and Daniel A Low. Tracking lung tissue motion and expansion/compression with inverse consistent image registration and spirometry. *Medical physics*, 34(6Part1):2155–2163, 2007. 2, 5

[14] Adrian V Dalca, Guha Balakrishnan, John Guttag, and Mert R Sabuncu. Unsupervised learning for fast probabilistic diffeomorphic registration. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I*, pages 729–738. Springer, 2018. 1

[15] Adrian V Dalca, Guha Balakrishnan, John Guttag, and Mert R Sabuncu. Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces. *Medical image analysis*, 57:226–236, 2019. 1

[16] Bin Duan, Ming Zhong, and Yan Yan. Towards saner deep image registration. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 12459–12468, 2023. 2

[17] Paul Dupuis, Ulf Grenander, and Michael I Miller. Variational problems on flows of diffeomorphisms for image matching. *Quarterly of applied mathematics*, pages 587–600, 1998. 3

[18] Bruce Fischl. Freesurfer. *Neuroimage*, 62(2):774–781, 2012. 4

[19] Eleftherios Garyfallidis, Matthew Brett, Bagrat Amirbekian, Ariel Rokem, Stefan Van Der Walt, Maxime Descoteaux, Ian Nimmo-Smith, and Dipy Contributors. Dipy, a library for the analysis of diffusion mri data. *Frontiers in neuroinformatics*, 8:8, 2014. 6

[20] Morteza Ghahremani, Mohammad Khateri, Bailiang Jian, Benedikt Wiestler, Ehsan Adeli, and Christian Wachinger. H-vit: A hierarchical vision transformer for deformable image registration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11513–11523, 2024. 3, 6

[21] Kun Han, Shanlin Sun, Xiangyi Yan, Chenyu You, Hao Tang, Junayed Naushad, Haoyu Ma, Deying Kong, and Xiaohui Xie. Diffeomorphic image registration with neural velocity field. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 1869–1879, 2023. 1, 3, 4

[22] Monica Hernandez, Matias N Bossa, and Salvador Olmos. Registration of anatomical images using geodesic paths of diffeomorphisms parameterized with stationary vector fields. In *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8. IEEE, 2007. 1

[23] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020. 1, 2, 4

[24] Malte Hoffmann, Benjamin Billot, Douglas N Greve, Juan Eugenio Iglesias, Bruce Fischl, and Adrian V Dalca. Synthmorph: learning contrast-invariant registration without acquired images. *IEEE transactions on medical imaging*, 41 (3):543–558, 2021. 1

[25] Ankita Joshi and Yi Hong. R2net: Efficient and flexible diffeomorphic image registration using lipschitz continuous residual networks. *Medical Image Analysis*, 89:102917, 2023. 6

[26] David N Kennedy, Christian Haselgrove, Steven M Hodge, Pallavi S Rane, Nikos Makris, and Jean A Frazier. Candishare: a resource for pediatric neuroimaging data. *Neuroinformatics*, 10:319–322, 2012. 5, 2

[27] Boah Kim, Dong Hwan Kim, Seong Ho Park, Jieun Kim, June-Goo Lee, and Jong Chul Ye. Cyclemorph: cycle consistent unsupervised deformable image registration. *Medical image analysis*, 71:102036, 2021. 2, 5

[28] Boah Kim, Inhwa Han, and Jong Chul Ye. Diffusemorph: Unsupervised deformable image registration using diffusion model. In *European conference on computer vision*, pages 347–364. Springer, 2022. 1, 3, 6

[29] Arno Klein and Jason Tourville. 101 labeled brain images and a consistent human cortical labeling protocol. *Frontiers in neuroscience*, 6:171, 2012. 5, 3

[30] Daniel S Marcus, Tracy H Wang, Jamie Parker, John G Csernansky, John C Morris, and Randy L Buckner. Open access series of imaging studies (oasis): cross-sectional mri data in young, middle aged, nondemented, and demented older adults. *Journal of cognitive neuroscience*, 19(9):1498–1507, 2007. 5, 2

[31] Mingyuan Meng, Dagan Feng, Lei Bi, and Jinman Kim. Correlation-aware coarse-to-fine mlps for deformable medical image registration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9645–9654, 2024. 3, 6

[32] Michael I Miller, Alain Trouvé, and Laurent Younes. On the metrics and euler-lagrange equations of computational anatomy. *Annual review of biomedical engineering*, 4(1): 375–405, 2002. 3

[33] Marc Modat, Gerard R Ridgway, Zeike A Taylor, Manja Lehmann, Josephine Barnes, David J Hawkes, Nick C Fox, and Sébastien Ourselin. Fast free-form deformation using graphics processing units. *Computer methods and programs in biomedicine*, 98(3):278–284, 2010. 1, 2

[34] Tony CW Mok and Albert Chung. Fast symmetric diffeomorphic image registration with convolutional neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4644–4653, 2020. 1, 3, 4, 5

[35] Tony CW Mok and Albert CS 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. 3, 5

[36] Tony CW Mok and Albert CS Chung. Conditional deformable image registration with convolutional neural network. 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 35–45. Springer, 2021. 1, 4

[37] William Peebles and Saining Xie. Scalable diffusion models with transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 4195–4205, 2023. 1, 2, 4

[38] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10684–10695, 2022. 1

[39] Daniel Rueckert, Luke I Sonoda, Carmel Hayes, Derek LG Hill, Martin O Leach, and David J Hawkes. Nonrigid registration using free-form deformations: application to breast mr images. *IEEE transactions on medical imaging*, 18(8): 712–721, 1999. 1

[40] Christos Sakaridis, Dengxin Dai, and Luc Van Gool. Acdc: The adverse conditions dataset with correspondences for semantic driving scene understanding. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 10765–10775, 2021. 5, 3

[41] David W Shattuck, Mubeena Mirza, Vitria Adisetiyo, Cornelius Hojatkashani, Georges Salamon, Katherine L Narr, Russell A Poldrack, Robert M Bilder, and Arthur W Toga. Construction of a 3d probabilistic atlas of human cortical structures. *Neuroimage*, 39(3):1064–1080, 2008. 5, 2

[42] Ameneh Sheikhjafari, Michelle Noga, Kumaradevan Punithakumar, and Nilanjan Ray. Unsupervised deformable image registration with fully connected generative neural network. In *Medical imaging with deep learning*, 2022. 1

[43] Zhengyang Shen, Xu Han, Zhenlin Xu, and Marc Niethammer. Networks for joint affine and non-parametric image registration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4224–4233, 2019. 1

[44] Leonard Siegert, Paul Fischer, Mattias P Heinrich, and Christian F Baumgartner. Pulpo: Probabilistic unsupervised laplacian pyramid registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 717–727. Springer, 2024. 3, 5

[45] Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International conference on machine learning*, pages 2256–2265. PMLR, 2015. 1, 2, 4

[46] Aristeidis Sotiras, Christos Davatzikos, and Nikos Paragios. Deformable medical image registration: A survey. *IEEE transactions on medical imaging*, 32(7):1153–1190, 2013. 1

[47] Shanlin Sun, Kun Han, Hao Tang, Deying Kong, Junayed Naushad, Xiangyi Yan, and Xiaohui Xie. Medical image registration via neural fields. *arXiv preprint arXiv:2206.03111*, 2022. 1

[48] J-P Thirion. Non-rigid matching using demons. In *Proceedings CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 245–251. IEEE, 1996. 2

[49] Lin Tian, Hastings Greer, François-Xavier Vialard, Roland Kwitt, Raúl San José Estépar, Richard Jarrett Rushmore, Nikolaos Makris, Sylvain Bouix, and Marc Niethammer. Gradicon: Approximate diffeomorphisms via gradient inverse consistency. In *Proceedings of the IEEE/CVF Con-*

ference on Computer Vision and Pattern Recognition, pages 18084–18094, 2023. 5

[50] Lin Tian, Soumyadip Sengupta, Hastings Greer, Raúl San José Estépar, and Marc Niethammer. Nephi: Neural deformation fields for approximately diffeomorphic medical image registration. *arXiv preprint arXiv:2309.07322*, 2023. 3, 5

[51] Nicholas J Tustison, Brian B Avants, and James C Gee. Learning image-based spatial transformations via convolutional neural networks: A review. *Magnetic resonance imaging*, 64:142–153, 2019. 1

[52] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, and Nicholas Ayache. Symmetric log-domain diffeomorphic registration: A demons-based approach. In *International conference on medical image computing and computer-assisted intervention*, pages 754–761. Springer, 2008. 2

[53] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, and Nicholas Ayache. Diffeomorphic demons: Efficient non-parametric image registration. *NeuroImage*, 45(1):S61–S72, 2009. 2

[54] Jelmer M Wolterink, Jesse C Zwienenberg, and Christoph Brune. Implicit neural representations for deformable image registration. In *International Conference on Medical Imaging with Deep Learning*, pages 1349–1359. PMLR, 2022. 1, 3

[55] Yifan Wu, Tom Z Jiahao, Jiancong Wang, Paul A Yushkevich, M Ani Hsieh, and James C Gee. Nodeo: A neural ordinary differential equation based optimization framework for deformable image registration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 20804–20813, 2022. 1, 3, 6

[56] Xiaoran Zhang, John C Stendahl, Lawrence H Staib, Albert J Sinusas, Alex Wong, and James S Duncan. Adaptive correspondence scoring for unsupervised medical image registration. In *European Conference on Computer Vision*, pages 76–92. Springer, 2024. 6

[57] Yungeng Zhang, Yuru Pei, and Hongbin Zha. Learning dual transformer network for diffeomorphic registration. 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 129–138. Springer, 2021. 1

[58] Yongtai Zhuo and Yiqing Shen. Diffusereg: Denoising diffusion model for obtaining deformation fields in unsupervised deformable image registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 597–607. Springer, 2024. 1, 3, 6

[59] Jing Zou, Noémie Debroux, Lihao Liu, Jing Qin, Carola-Bibiane Schönlieb, and Angelica I Aviles-Rivero. Homeomorphic image registration via conformal-invariant hyperelastic regularisation. *arXiv preprint arXiv:2303.08113*, 2023. 1