



This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

The Right Prior for the Right Deformation: Rethinking Continuous Deformable Image Registration

Hengjie Liu¹[0000-0001-5890-2291], Chushu Shen^{2,3}[0009-0006-5865-4754], Dan Ruan^{3,4}[0000-0003-3400-7684], and Ke Sheng¹[0000-0002-6696-5409]

¹ Department of Radiation Oncology, University of California, San Francisco, San Francisco, CA, USA

{hengjie.liu, ke.sheng}@ucsf.edu

² Biomedical Imaging Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA

³ Department of Bioengineering, University of California, Los Angeles, Los Angeles, CA, USA

⁴ Department of Radiation Oncology, University of California, Los Angeles, Los Angeles, CA, USA

Abstract. Deformable image registration models implicitly encode deformation priors through their parametrization and optimization. In this work, we conduct a validation study on continuous registration methods to examine how these implicit priors affect performance across different registration tasks. Classic B-Spline transformations impose locality, smoothness, and scale through their control-point structure, whereas recent INR-based methods impose different priors through neural parameterization and optimization. We compare INR-Dense (IDIR), which directly models a dense displacement field using a SIREN-based INR; INR-BSCP (SINR), which predicts B-Spline control points with an INR; D-BSCP, which directly optimizes single-scale B-Spline control points; and MR-D-BSCP, which adds a multiresolution coarse-to-fine scheme. Experiments on inter-subject brain MR registration (OASIS) and intra-subject exhale-to-inhale lung CT registration (DIR-LAB 4DCT) reveal different behavior across deformation regimes. On OASIS, where deformations are moderate but locally complex, D-BSCP matches or slightly outperforms INR-BSCP, suggesting that the B-Spline parameterization accounts for much of INR-BSCP’s effectiveness. On DIR-LAB 4DCT, where respiratory motion is larger and more coherent, single-scale B-Spline methods (D-BSCP and INR-BSCP) are less suitable, while INR-Dense and MR-D-BSCP are more effective. Across both tasks, MR-D-BSCP achieves the best performance among the tested continuous parameterizations. These findings highlight that registration accuracy depends strongly on matching the induced deformation prior to the target motion pattern, and support prior-deformation matching as a practical design principle for medical image registration. Our code will be available at <https://github.com/HengjieLiu/RightPriorDIR>.

Keywords: Continuous Deformable Image Registration · Implicit Neural Representation · B-Spline · Validation.

1 Introduction

Deformable image registration (DIR) estimates a nonlinear spatial transformation that aligns corresponding anatomy across images. It is central to various medical image applications including longitudinal analysis, atlas construction, motion correction, image-guided intervention and so on. The transformation can be parameterized either as a continuous function or as a discrete field. Classical B-Spline free-form deformation (FFD) has been a successful continuous method because it reduces the degrees of freedom while flexibly representing diverse nonrigid deformations [22,14]. Discrete dense displacement fields, such as those optimized by Demons [26] and MRF-based methods [7], are also widely used.

In the past decade, discrete displacement fields parameterized by deep neural networks became popular due to strong performance and fast inference after population training [1,4,8]. However, learning-based methods can degrade in out-of-distribution settings, and pair-specific optimization remains the most robust general-purpose option [11,10]. Implicit neural representations (INRs) have recently renewed interest in continuous pair-specific registration [28,2,23,6]. Instead of storing deformation on a voxel grid, an INR represents it as a coordinate-conditioned function, typically a multilayer perceptron (MLP), that maps continuous spatial coordinates to displacement vectors. This enables arbitrary coordinate queries and analytic derivatives, naturally supporting gradient-based regularization. IDIR [28] first demonstrated this idea for exhale-to-inhale lung CT registration using a SIREN MLP [25]. SINR later argued that dense INR displacement prediction can be suboptimal for inter-subject brain MR registration, and instead used an INR to predict B-Spline FFD control-point values [23].

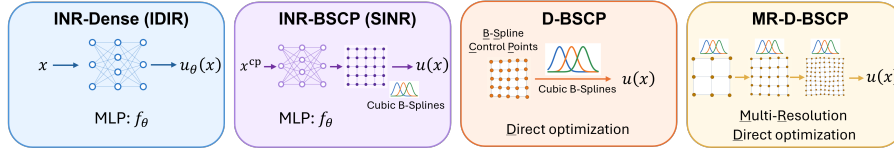
These results raise an often-overlooked question: when continuous registration improves, are the gains due to the INR, the spline parameterization, or a better match between the model and the deformation pattern? This matters because deformation patterns vary substantially across registration tasks. For example, inter-subject brain MR registration often involves moderate displacements with complex local variation, whereas exhale-to-inhale lung CT registration exhibits larger, more directionally coherent, and smoother respiratory motion. Thus, a continuous "off-grid" model effective in one setting may not generalize to another.

We propose a validation study of continuous DIR parameterizations from this prior-matching perspective, focusing on their accuracy-regularity spectrum [24]. For clarity, throughout the remainder of the paper, we use unified notation based on deformation parameterization: **INR-Dense** for IDIR, which directly models a continuous displacement field with a SIREN-based INR; **INR-BSCP** for SINR, which predicts B-Spline control points with an INR; **D-BSCP** for direct B-Spline control-point optimization; and **MR-D-BSCP** for its multiresolution coarse-to-fine extension. We compare these four core parameterizations on inter-subject

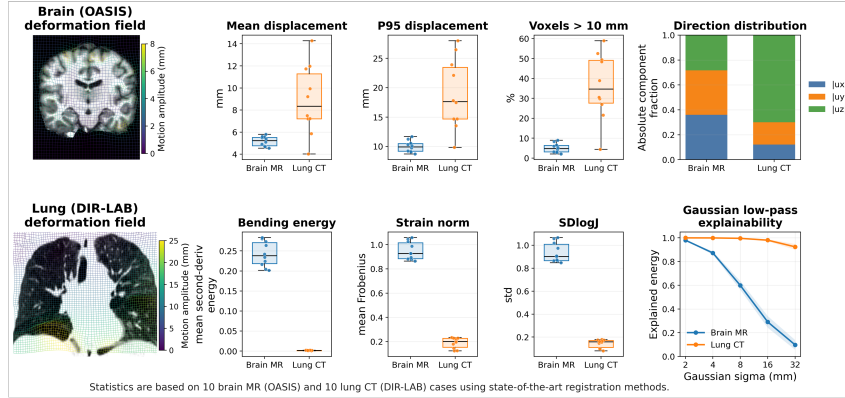
brain MR registration and intra-subject exhale-to-inhale lung CT registration (Fig. 1). We additionally evaluate the recently proposed Dual-INR [6], a multi-resolution INR approach for lung DIR, on the lung registration task and denote it **MR-INR-Dense**.

We find that D-BSCP can match INR-BSCP on brain MR registration, suggesting that the B-Spline parameterization explains much of INR-BSCP’s performance. For lung CT, however, single-scale B-Spline methods, including INR-BSCP and D-BSCP, are less suitable for large respiratory motion, whereas INR-Dense remains strong and MR-D-BSCP substantially improves the spline-based formulation. MR-INR-Dense also provides a modest improvement over INR-Dense. Overall, MR-D-BSCP achieves the best performance. Our analysis supports a simple but often overlooked principle: registration quality depends strongly on how well the model-induced deformation prior matches the target deformation pattern, including displacement range, smoothness, locality, directionality, frequency content, and optimization path. This also underscores the need for careful validation of new DIR methods to identify which components actually drive registration performance [13,16].

A. Continuous Registration Parameterizations



B. Different Deformation Patterns



Off-grid is not enough: model priors should match deformation characteristics.

Fig. 1. Overview of the tested continuous registration parameterizations and registration tasks.

2 Methods

2.1 Problem Formulation for DIR

DIR estimates a displacement field $\mathbf{u}(\mathbf{x})$ that aligns a moving image M to a fixed image F by minimizing the following objective function:

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u}} [\mathcal{L}_{\text{sim}}(M \circ \phi, F) + \lambda \mathcal{R}(\mathbf{u})], \quad (1)$$

where the spatial transformation is given by $\phi(\mathbf{x}) = \mathbf{x} + \mathbf{u}(\mathbf{x})$. Here, $\mathcal{L}_{\text{sim}}$ denotes the image-similarity loss, $\mathcal{R}$ denotes the regularization term, and λ controls the trade-off between similarity and regularity. Following the alignment-regularity evaluation framework of [24], we compare methods across regularization strengths to enable fair comparison and identify Pareto fronts.

2.2 Tested Continuous Registration Models

Continuous registration methods represent the displacement field $\mathbf{u} : \Omega \rightarrow \mathbb{R}^3$ as a continuous function over the image domain Ω .

B-Spline FFD. It parameterizes $\mathbf{u}(\mathbf{x})$ by a regular lattice of control-point displacements. For a spatial location $\mathbf{x}$, the displacement is computed by tensor-product cubic B-Spline interpolation:

$$\mathbf{u}(\mathbf{x}) = \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 B_l(\xi) B_m(\eta) B_n(\zeta) \mathbf{c}_{i+l, j+m, k+n}, \quad (2)$$

where $\mathbf{c}_{i+l, j+m, k+n} \in \mathbb{R}^3$ denotes a neighboring control-point displacement, $(\xi, \eta, \zeta) \in [0, 1]^3$ are the normalized local coordinates of $\mathbf{x}$ within the enclosing lattice cell, and B_l , B_m , and B_n are cubic B-Spline basis functions. Due to compact support, each displacement depends only on a local control-point neighborhood, and each control point affects only a finite spatial region. Thus, the control-point spacing sets the deformation scale: coarser grids favor smoother, more global deformations, whereas finer grids allow more localized motion.

We evaluate both single-scale direct B-Spline control-point optimization (**D-BSCP**) and its multiresolution variant (**MR-D-BSCP**). D-BSCP optimizes control-point displacements at a fixed spacing, e.g., 4 voxels or mm. MR-D-BSCP uses a coarse-to-fine schedule, e.g., 32-16-8-4, with a matching image pyramid in which images are smoothed and downsampled at coarser stages. Each stage initializes the next finer lattice from the previous stage. While general spacing schedules require field upsampling via refitting [14], our dyadic $2 \times 2 \times 2$ schedule yields nested B-Spline lattices, enabling fast analytic knot insertion.

INR-Dense (IDIR). $\mathbf{u}$ is parameterized with a coordinate-based SIREN MLP, f_θ , which takes any spatial coordinate $\mathbf{x} \in \Omega$ as input and directly predicts the displacement:

$$\mathbf{u}_\theta(\mathbf{x}) = f_\theta(\mathbf{x}), \quad \phi_\theta(\mathbf{x}) = \mathbf{x} + \mathbf{u}_\theta(\mathbf{x}). \quad (3)$$

The network is optimized for each image pair. Because f_θ is differentiable with respect to its input coordinates, spatial derivatives of the transformation can be computed by automatic differentiation.

INR-BSCP (SINR). It uses the same coordinate-based SIREN MLP but predicts B-Spline control-point displacements instead of dense displacements:

$$\mathbf{c}_{p,q,r} = f_\theta(\mathbf{x}_{p,q,r}^{\text{cp}}), \quad (4)$$

where $\mathbf{x}_{p,q,r}^{\text{cp}}$ is the coordinate of a B-Spline control point. The final displacement field is then obtained by Eq. 2.

3 Experiments

3.1 Datasets and Baselines

We evaluated two registration tasks: inter-subject brain MR registration on OASIS [18] and intra-subject exhale-to-inhale lung CT registration on DIR-LAB 4DCT [3]. OASIS contains 414 1mm isotropic T1-weighted MR images. We used the 314/20/80 train/validation/test split for comparison with learning-based baselines and evaluated 100 randomly sampled test pairs. Baselines included four learning-based methods: VoxelMorph [1], TransMorph [4], VFA [17], and SITReg [9], with VFA and SITReg among the top LUMIR challenge benchmarks [5]. We also included Greedy [29,12] as a strong optimization-based baseline.

For DIR-LAB 4DCT, we evaluated all ten 4DCT cases using the two extreme respiratory phases (T00 and T50). We followed the convention of pTVReg [27], a state-of-the-art optimization-based method for this task: scans were resampled to 1mm isotropic spacing for registration, while landmark TRE was computed after resampling the displacement field back to the original image grid. Since this preprocessing differs from the original INR-Dense setting, we reran INR-Dense on the isotropic images. Given the small cohort of only ten registration pairs, we did not evaluate learning-based baselines on this task.

3.2 Implementation Details

We used the official IDIR and SINR codebases to implement INR-Dense and INR-BSCP, respectively, keeping their SIREN architectures, sampling strategies, image-similarity losses, and optimizer settings unchanged. Both use three hidden layers with 256 units per layer. For MR-INR-Dense, we adapted the official Dual-INR implementation and evaluated it on DIR-LAB 4DCT, removing only the

lobe-segmentation loss, which we found had negligible effect on performance. We implemented D-BSCP and MR-D-BSCP in PyTorch to directly optimize B-Spline control-point displacements, isolating the effect of the INR from the B-Spline control-point parameterization.

All methods were optimized with Adam for a total budget of 2,500 steps to ensure convergence, with MR-D-BSCP steps divided across resolution levels. Because INR-based and direct B-Spline models have different parameter scales and conditioning, we did not enforce a shared learning rate. Instead, we applied the same coarse sweep over $\{10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$ to both families and selected the best stable setting for each. For INR-based methods, 10^{-2} and 10^{-3} were unstable, while the official 10^{-4} setting was stable and therefore selected. For direct B-Spline methods, we used 10^{-3} , which was stable and gave the best final performance, while 10^{-2} converged faster but slightly reduced accuracy. All experiments were run on NVIDIA RTX 6000 Ada Generation GPUs with 48 GB of memory.

3.3 Evaluation

Rather than comparing methods at a single selected hyperparameter, we used the alignment-regularity characteristic (ARC) framework for a fairer comparison across methods [24]. ARC compares the full alignment-regularity trade-off spectrum, avoiding conclusions based on a single operating point where better alignment may come at the cost of deformation regularity. For OASIS, alignment was measured by the mean Dice score and mean 95th percentile Hausdorff distance (HD95) over 35 segmentation labels. For DIR-LAB 4DCT, it was measured by landmark target registration error (TRE) over 300 manually annotated landmark pairs per case. Regularity was measured within foreground masks using the standard deviation of the log-Jacobian determinant ($\text{SD}(\log J)$). Fixed external baselines (four learning-based methods and Greedy) were included as single operating points.

We further performed a target-field fitting diagnostic to separate model representation capacity from registration capacity. Each parameterization was trained with a supervised mean-squared error (MSE) loss to fit a strong external reference displacement field: SITReg for OASIS and pTVReg for DIR-LAB 4DCT, both selected from high-performing Pareto-front methods. These fields were not treated as ground truth, but as plausible high-performing deformations for testing how well each parameterization can represent the target when given directly. We ran supervised fitting on 10 random pairs from OASIS and all 10 pairs from DIR-LAB 4DCT, and compared with the unsupervised registration optimization. For unsupervised registration, we selected regularization weights based on the Pareto sweeps so that the compared methods operated at approximately matched deformation regularity, measured by foreground $\text{SD}(\log J)$.

4 Results

4.1 Registration Results

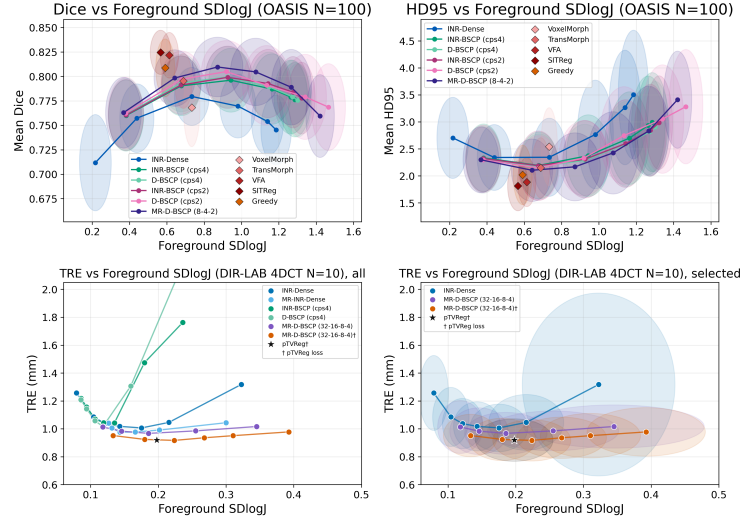


Fig. 2. Registration performance across the alignment-regularity spectrum on OASIS brain MR (top) and DIR-LAB 4DCT (bottom). Ellipse axes indicate standard deviation across cases.

Fig. 2 summarizes registration performance across the alignment-regularity spectrum. On OASIS, D-BSCP closely follows INR-BSCP across the Dice versus $SD(\log J)$ and HD95 versus $SD(\log J)$ curves at both 2 mm and 4 mm control-point spacing (cps). In several regions, D-BSCP performs slightly better. Thus, similar performance to INR-BSCP can be obtained by directly optimizing the B-Spline control points without the INR. MR-D-BSCP (8-4-2) achieves the best performance among the tested continuous parameterizations. In contrast, INR-Dense performs worse and occupies a smoother region of the trade-off, suggesting that its model prior may limit its ability to capture locally varying inter-subject brain deformations. Among external references, SITReg, VFA, and Greedy remain stronger operating points, while earlier learning-based methods without registration-specific design, especially multiresolution refinement, are less competitive [13,16].

DIR-LAB 4DCT shows a different behavior. Single-scale INR-BSCP and D-BSCP are less robust to large respiratory motion and can fail sharply at weaker regularization. INR-Dense is the most robust single-scale method on this task. MR-D-BSCP (32-16-8-4) substantially improves the spline-based formulation and closely matches the state-of-the-art pTVReg performance when using

Table 1. Runtime-reduced OASIS schedules versus their 2500-step counterparts. Accuracy is OASIS-100. Runtime is input-image to saved dense displacement field, including optimization and DVF output, excluding metric evaluation. Cold indicates a fresh process/CUDA context; warm indicates persistent-process timing after warmup.

Method	Dice	Steps	Cold (s)		Warm (s)		Peak GPU (GB)
	base→fast	base→fast	base→fast	base→fast	base→fast	base→fast	
MR-D-BSCP (8-4-2)	0.810→0.809	2500→440	152.0→101.1	62.4→11.0			3.93
D-BSCP (cps2)	0.805→0.805	2500→431	135.4→76.9	80.9→14.4			2.16
D-BSCP (cps4)	0.799→0.797	2500→477	126.3→78.6	80.9→15.3			1.83
INR-BSCP (cps2)	0.799→0.793	2500→799	370.9→171.4	277.9→90.2			7.72
INR-BSCP (cps4)	0.796→0.786	2500→644	175.0→90.4	107.8→29.2			2.58

the same loss function as pTVReg (0.917 ± 0.148 mm for MR-D-BSCP versus 0.919 ± 0.152 mm for pTVReg). These results suggest that large respiratory motion favors either a dense INR prior that promotes smooth, coherent deformation or, for local B-Spline control points, a coarse-to-fine optimization path that captures large displacement before local refinement.

MR-INR-Dense, adapted from Dual-INR [6], further supports the benefits of multiscale modeling for lung CT by decomposing motion into coarse and fine INR branches. It improves over INR-Dense, although it still does not match pTVReg’s TRE. However, unlike B-Spline control lattices, INR parameters cannot be directly upsampled for finer-resolution refinement and instead require residual branches, composition, or continued optimization.

We also evaluated whether these pair-specific optimizers can be made faster without sacrificing the registration accuracy. We focus this runtime analysis on OASIS, where all registrations share the same image size; DIR-LAB cases differ in crop size and motion difficulty, making a compact runtime comparison less controlled. For each spline-based method, we kept the objective and regularization weight fixed and screened learning-rate multipliers, shortened optimization-step budgets, and adaptive early stopping on OASIS. The selected schedules were then evaluated on the OASIS-100 test set and timed from input images to the saved dense displacement field, including full optimization but excluding evaluation metrics. Table 1 reports both cold-process and warm-process runtimes, since the first case in a Python/CUDA process includes one-time library and kernel setup. The optimized schedules preserve OASIS-100 Dice closely while substantially reducing runtime; MR-D-BSCP in particular remains the most accurate tested continuous parameterization while reducing warm input-to-DVF time from about 62 s to 11 s.

4.2 Target-field Fitting vs Registration Diagnostics

Fig. 3 examines representation capacity versus registration optimization behavior by comparing supervised target-field fitting with unsupervised registration. On OASIS, the 10 sampled pairs show highly consistent behavior across methods, so we report only the population-level results. DIR-LAB is more heterogeneous.

We therefore additionally show case01, where the methods behave similarly well, and case08, where single-scale methods struggle under large motion.

Under supervised fitting, D-BSCP fits the reference fields well, with MR-D-BSCP achieving the lowest fitting errors. INR-BSCP and INR-Dense generally show higher errors. This pattern is consistent across the OASIS and DIR-LAB cases.

The contrast is most evident on the large-motion DIR-LAB case08. D-BSCP fits the target field well under supervised fitting, showing that the single-scale B-Spline model has sufficient representation capacity, yet it fails during unsupervised registration. INR-Dense shows the opposite trend. Its target-field fitting error is substantially higher, but it achieves better registration under large motion. When the model prior matches the deformation pattern, as in INR-Dense for smooth and coherent lung motion, good registration can still be achieved despite lower representation capability. Conversely, when the prior or optimization path is mismatched, high representation capacity alone may not prevent failure. In this case, the strong improvement from multiresolution optimization supports the interpretation that the single-scale methods fail because they converge to poor local minima. These results suggest that both the parameterization and the optimization scheme should be chosen according to the deformation properties of the target task.

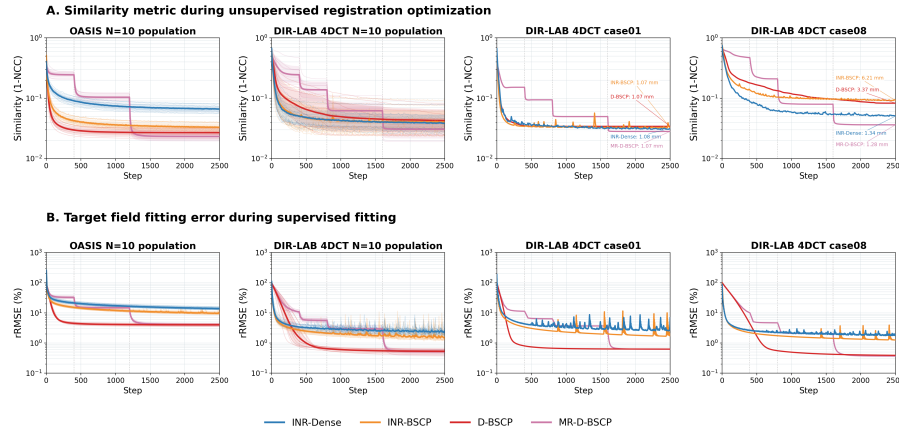


Fig. 3. Unsupervised registration versus supervised target-field fitting diagnostic for population and representative cases. In population panels, faint lines are individual cases, thick lines are log-space population means, and shaded regions show ± 1 log-space standard deviation.

5 Conclusion and Discussion

We presented a controlled validation study of continuous deformable registration parameterizations across inter-subject brain MR and exhale-to-inhale lung CT registration. Our results show that the success of an "off-grid" registration model is not determined by the use of a better continuous model, but by whether the induced deformation prior matches the target motion pattern. On OASIS, direct B-Spline control-point (D-BSCP) optimization closely matches INR-BSCP (SINR), suggesting that the B-Spline control-point prior explains much of INR-BSCP's favorable behavior. Moreover, our runtime-reduction study shows that D-BSCP and MR-D-BSCP can be made substantially faster on OASIS while preserving registration accuracy. On DIR-LAB, large respiratory motion requires long-range capture behavior, where INR-Dense (IDIR) and especially multiresolution methods are more suitable than single-scale B-Spline methods.

Our findings echo recent studies in deep learning-based DIR, which show that registration-specific designs matter more than naively adopting a new network architecture [13,16]. Our study extends this message to continuous pair-specific registration: INR-based methods are useful, but they are not automatically superior to well-established B-Spline parameterizations when validated carefully. This motivates more rigorous validation of new registration methods, including ablation of the components that actually drive performance. We also highlight the importance of evaluating methods across the alignment-regularity spectrum [24], rather than at a single operating point, to support fair comparison across different methods.

At the same time, our study does not isolate all sources of prior, particularly the effect of different regularization forms, and broader validation across registration tasks is still needed. More broadly, these results suggest that registration models should be designed around the deformation properties of the target anatomy. Similar ideas have been explored through sliding-motion, anatomy-adaptive, and biomechanics-informed priors [19,15,20]. Developing task-specific deformation priors remains a promising direction for DIR [21].

Acknowledgments. This study was funded by NIH R01EB031577.

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

References

1. Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J., Dalca, A.V.: VoxelMorph: A learning framework for deformable medical image registration. *IEEE Transactions on Medical Imaging* **38**(8), 1788–1800 (2019). <https://doi.org/10.1109/TMI.2019.2897538>
2. Byra, M., Poon, C., Rachmadi, M.F., Schlachter, M., Skibbe, H.: Exploring the performance of implicit neural representations for brain image registration. *Scientific Reports* **13**(1), 17334 (2023). <https://doi.org/10.1038/s41598-023-44517-5>

3. Castillo, R., Castillo, E., Guerra, R., et al.: A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. *Physics in Medicine and Biology* **54**(7), 1849–1870 (2009). <https://doi.org/10.1088/0031-9155/54/7/001>
4. Chen, J., Frey, E.C., He, Y., et al.: TransMorph: Transformer for unsupervised medical image registration. *Medical Image Analysis* **82**, 102615 (2022). <https://doi.org/10.1016/j.media.2022.102615>
5. Chen, J., Wei, S., Honkamaa, J., et al.: Beyond the LUMIR challenge: The pathway to foundational registration models. *Medical Image Analysis* **113**, 104175 (2026). <https://doi.org/10.1016/j.media.2026.104175>
6. Gebauer, J.B., Nielsen, M., Madesta, F., Werner, R., Sentker, T.: Robust multi-scale implicit neural representations for large-deformation lung registration. In: *Proceedings of The 9th International Conference on Medical Imaging with Deep Learning. Proceedings of Machine Learning Research*, vol. 315, pp. 3089–3102. PMLR (2026)
7. Heinrich, M.P., Jenkinson, M., Brady, M., Schnabel, J.A.: MRF-based deformable registration and ventilation estimation of lung CT. *IEEE Transactions on Medical Imaging* **32**(7), 1239–1248 (2013). <https://doi.org/10.1109/TMI.2013.2246577>
8. Hering, A., Hansen, L., Mok, T.C.W., et al.: Learn2Reg: Comprehensive multi-task medical image registration challenge, dataset and evaluation in the era of deep learning. *IEEE Transactions on Medical Imaging* **42**(3), 697–712 (2023). <https://doi.org/10.1109/TMI.2022.3213983>
9. Honkamaa, J., Marttinen, P.: SITReg: Multi-resolution architecture for symmetric, inverse consistent, and topology preserving image registration. *Machine Learning for Biomedical Imaging* **2**, 2148–2194 (2024). <https://doi.org/10.59275/j.melba.2024-276b>
10. Jena, R., Chaudhari, P., Gee, J.: The LUMirage: An independent evaluation of zero-shot performance in the LUMIR challenge. In: *Proceedings of The 9th International Conference on Medical Imaging with Deep Learning. Proceedings of Machine Learning Research*, vol. 315, pp. 4134–4165. PMLR (2026)
11. Jena, R., Chaudhari, P., Gee, J.C.: Adaptive riemannian optimization for multi-scale diffeomorphic matching. *Nature Communications* **17**(1), 4774 (2026). <https://doi.org/10.1038/s41467-026-72508-3>
12. Jena, R., Sethi, D., Chaudhari, P., Gee, J.C.: Deep learning in medical image registration: Magic or mirage? In: *Advances in Neural Information Processing Systems* **37**, pp. 108331–108353. Neural Information Processing Systems Foundation, Inc. (2024). <https://doi.org/10.52202/079017-3439>
13. Jian, B., Pan, J., Ghahremani, M., et al.: Mamba? catch the hype or rethink what really helps for image registration. In: *Biomedical Image Registration. Lecture Notes in Computer Science*, vol. 15249, pp. 86–97. Springer Nature Switzerland (2024). https://doi.org/10.1007/978-3-031-73480-9_7
14. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.W.: elastix: A toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging* **29**(1), 196–205 (2010). <https://doi.org/10.1109/TMI.2009.2035616>
15. Liu, H., McKenzie, E., Xu, D., et al.: MUsculo-Skeleton-Aware (MUSA) deep learning for anatomically guided head-and-neck CT deformable registration. *Medical Image Analysis* **99**, 103351 (2025). <https://doi.org/10.1016/j.media.2024.103351>
16. Liu, H., Ruan, D., Sheng, K.: Unsupervised deformable image registration revisited: Enhancing performance with registration-specific designs. In: *Medical Imaging with Deep Learning* (2025), mIDL 2025 Short Paper

17. Liu, Y., Chen, J., Zuo, L., Carass, A., Prince, J.L.: Vector field attention for deformable image registration. *Journal of Medical Imaging* **11**(6), 064001 (2024). <https://doi.org/10.1117/1.JMI.11.6.064001>
18. Marcus, D.S., Wang, T.H., Parker, J., et al.: 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). <https://doi.org/10.1162/jocn.2007.19.9.1498>
19. Papież, B.W., Heinrich, M.P., Fehrenbach, J., Risser, L., Schnabel, J.A.: An implicit sliding-motion preserving regularisation via bilateral filtering for deformable image registration. *Medical Image Analysis* **18**(8), 1299–1311 (2014). <https://doi.org/10.1016/j.media.2014.05.005>
20. Qin, C., Wang, S., Chen, C., Bai, W., Rueckert, D.: Generative myocardial motion tracking via latent space exploration with biomechanics-informed prior. *Medical Image Analysis* **83**, 102682 (2023). <https://doi.org/10.1016/j.media.2022.102682>
21. Reithmeir, A., Spieker, V., Sideri-Lampretsa, V., et al.: From model based to learned regularization in medical image registration: A comprehensive review. *Medical Image Analysis* **108**, 103854 (2026). <https://doi.org/10.1016/j.media.2025.103854>
22. Rueckert, D., Sonoda, L.I., Hayes, C., et al.: Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Transactions on Medical Imaging* **18**(8), 712–721 (1999). <https://doi.org/10.1109/42.796284>
23. Sideri-Lampretsa, V., McGinnis, J., Qiu, H., et al.: SINR: Spline-enhanced implicit neural representation for multi-modal registration. In: *Proceedings of The 7th International Conference on Medical Imaging with Deep Learning. Proceedings of Machine Learning Research*, vol. 250, pp. 1462–1474. PMLR (2024)
24. Sideri-Lampretsa, V., Rueckert, D., Qiu, H.: Evaluation of deformable image registration under alignment-regularity trade-off. In: *Bridging Regulatory Science and Medical Imaging Evaluation; and Distributed, Collaborative, and Federated Learning. Lecture Notes in Computer Science*, vol. 16135, pp. 5–14. Springer, Cham (2026). https://doi.org/10.1007/978-3-032-05663-4_1
25. Sitzmann, V., Martel, J.N.P., Bergman, A.W., Lindell, D.B., Wetzstein, G.: Implicit neural representations with periodic activation functions. In: *Advances in Neural Information Processing Systems* 33. vol. 33, pp. 7462–7473 (2020)
26. Thirion, J.P.: Image matching as a diffusion process: an analogy with Maxwell’s demons. *Medical Image Analysis* **2**(3), 243–260 (1998). [https://doi.org/10.1016/S1361-8415\(98\)80022-4](https://doi.org/10.1016/S1361-8415(98)80022-4)
27. Vishnevskiy, V., Gass, T., Szekely, G., Tanner, C., Goksel, O.: Isotropic total variation regularization of displacements in parametric image registration. *IEEE Transactions on Medical Imaging* **36**(2), 385–395 (2017). <https://doi.org/10.1109/TMI.2016.2610583>
28. Wolterink, J.M., Zwienenberg, J.C., Brune, C.: Implicit neural representations for deformable image registration. In: *Proceedings of The 5th International Conference on Medical Imaging with Deep Learning. Proceedings of Machine Learning Research*, vol. 172, pp. 1349–1359. PMLR (2022)
29. Yushkevich, P.A., Pluta, J., Wang, H., et al.: IC-P-174: Fast automatic segmentation of hippocampal subfields and medial temporal lobe subregions in 3 tesla and 7 tesla T2-weighted MRI. *Alzheimer’s & Dementia* **12**, P126–P127 (2016). <https://doi.org/10.1016/j.jalz.2016.06.205>