Over-parameterising of optimisation for sparse 3D medical image registration

Mattias P. Heinrich[1[0000−0002−7489−1972]]

Institute of Medical Informatics, University of Luebeck, Germany
mattias.heinrich@uni-luebeck.de

Abstract. Despite advances in self-supervised learning for robust end-to-end medical image registration models, instance-based optimisation for the alignment of geometric or semantic features still remains highly effective for complex clinical tasks. Recently, the fitting of SIREN-based neural fields has challenged the direct optimisation of transformation parameters. In this work, we provide insights into these underlying mechanisms by drawing comparisons to explicit strategies for over-parameterisation. By evaluating a unified explicit over-parameterisation scheme that combines 12-degree-of-freedom (DoF) log-Euclidean polyaffine matrix prediction with spatially adaptive multi-scale smoothing, we empirically demonstrate a substantial positive impact on first-order optimisers (Adam), whereas second-order methods (L-BFGS) remain vulnerable to local minima in sparse formulations. The SIREN-based neural field in contrast benefits less, because it inherently captures local coarse-to-fine smoothness as well as global rotation, shears, or scaling. A symmetric objective aligning multi-scale MIND features in a mutual mid-point space enforces inverse consistency throughout optimisation. Combining these ideas sets new state-of-the-art performance for unsupervised inter-subject abdominal CT registration and improves upon previous continuous-optimisation-based approaches for large-deformation COPD inspiration-expiration alignment.

Keywords: Sparse Deformable 3D Registration · SIREN-based neural fields · Over-parameterisation · Multi-scale Blending · Polyaffine

1 Introduction

Image registration is a fundamental task in medical image analysis, enabling the alignment of images from different time points, modalities, or subjects. The research field has seen significant advancements with the advent of deep learning, particularly in self-supervised learning approaches that learn robust end-to-end registration models [26,7,22]. However, optimisation-based methods [2,11,18], which directly optimise transformation parameters to align geometric or semantic features, continue to demonstrate highly competitive performance in many clinical applications. Recently, the fitting of SIREN-based neural fields [27] has emerged as a promising alternative to traditional optimisation techniques. These neural fields leverage neural networks to model complex deformations, capturing both local and global transformations through a concept of lifting [24]. By expanding the representation capacity to map a low-DoF coordinate space into a higher-dimensional parameter space, a process termed *lifting* [24], the non-convex optimisation landscape can be smoothed. That means one indirectly fits a continuous function that maps static inputs (e.g. coordinates) to instance based deformation parameters. This can be further stabilised by symmetry contraints [8], diffeomorphisms [6] or B-spline models [17]. Despite their potential, the performance of SIREN-based neural fields can be limited by the choice of network architecture and the inherent challenges of optimising high-dimensional parameter spaces, leading to hybrid approaches that aim to predict good starting points for neural field fine-tuning or instance optimisation [16,23,5,9].

To the best of our knowledge, the impact of explicit and implicit over-parameterisation on instance-based registration has not been systematically studied. Proven explicit formulations such as log-Euclidean polyaffine transforms [1,21] or spatially varying smoothness [19,15] have not been put in direct context to implicit neural fields. Hence, a fair comprehensive comparison of different choices of lifting the underlying optimisation in a higher-dimensional space to ease convergence is so far missing.

In this work, we explore the effects of explicit over-parameterisation in two distinct ways: 1) by predicting polyaffine log-Euclidean matrices with 12 parameters instead of the traditional 3 parameters for each control point, and 2) by optimising a spatially varying deformation field smoothness. We hypothesise that these strategies can improve convergence and registration quality, particularly in scenarios with poor initial alignment and complex, non-linear deformations. All methods and experiments are designed without relying on regular grids but rather any sparse set of (off-grid) control points, which improve computational efficiency and allow for more flexible and adaptive registration strategies.

Source code with further implementation details is available at `https://github.com/mattiaspaul/overparam_reg`.

2 Methods

To investigate alternative schemes of over-parameterisation, we pose image registration as a sparse control-based 3D alignment of features via instance optimisation coupled with dense interpolation strategies. Given two images, a fixed volume $\mathbf{I}_f$ and a moving volume $\mathbf{I}_m$, we first extract multi-scale geometric Modality Independent Neighbourhood Descriptor (MIND-SSC) features [12] on a dense voxel matrix grid $\Omega \subset \mathbb{R}^3$ of dimensions $H \times W \times D$ using a fixed descriptor function Ψ. Next, we define a set of K sparse, off-grid control points $\mathcal{P} = \{\mathbf{p}_k\}_{k=1}^{K} \subset \mathbb{R}^3$. While these control points can adaptively leverage anatomically salient landmarks derived from semantic point clouds, in this work $\mathcal{P}$ is instantiated as a regular control-point lattice ($11 \times 11 \times 11 = 1{,}331$ points for COPDgene and $17 \times 17 \times 17 = 4{,}913$ points for AMOS22) to systematically evaluate spatial coverage. Akin to [2], our objective is to optimise a pair of

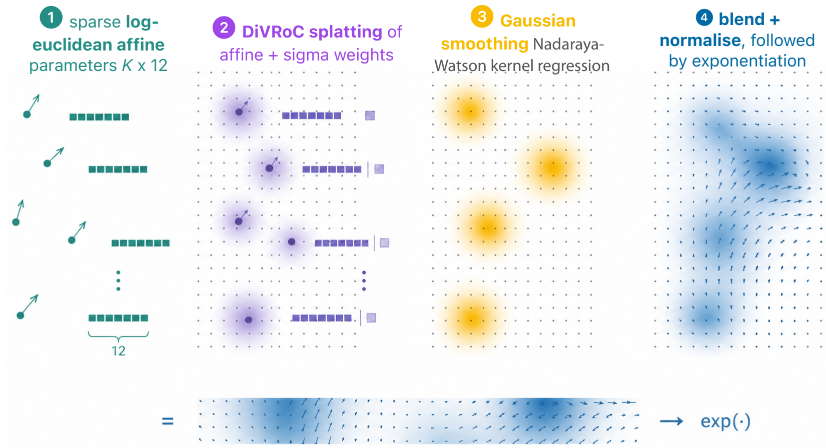


Fig. 1: Overview of the proposed off-grid, multi-scale log-Euclidean polyaffine transformation model. **Step 1 (Sparse Parameterization):** The continuous space is parameterised by K unstructured control points $\mathcal{P}$, each predicting a 12-DoF log-Euclidean matrix $\mathbf{M}_k \in \mathfrak{aff}(3)$ and multi-scale softmax probabilities $\pi_k \in \Delta^{M-1}$. **Step 2 (Differentiable Volumetric Rasterisation):** Sparse parameters, scale probabilities, and localised density indicators are projected onto the discrete voxel grid Ω using highly efficient DiVRoC splatting to yield the voxelised sparse fields ($\mathbf{A}_{\text{sparse}}$, $\mathbf{\Pi}_{\text{sparse},m}$, $\mathbf{W}_{\text{sparse}}$). **Step 3 (Multi-Scale Spatially Adaptive Regression):** These sparse grid fields are convolved across a spectrum of discrete 3D separable Gaussian kernels G_{σ_m}, establishing scale-specific intermediate representations via Nadaraya-Watson kernel regression. **Step 4 (Dense Blending and Matrix Field Integration):** Voxel-wise scale probabilities are renormalised to form a convex combination that produces the final continuous dense log-affine matrix field $\mathbf{A}_{\text{dense}}(\mathbf{x})$. This field is integrated via voxel-wise matrix exponentials $\mathbf{T}_{1,2}(\mathbf{x}) = \exp(\pm 0.5\mathbf{A}_{\text{dense}}(\mathbf{x}))$ to yield invertible midpoint transformations.

dense, symmetric, diffeomorphic mappings, forward (ϕ_1) and backward (ϕ_2), to a mid-point space that minimises structural feature dissimilarity in a mutual coordinate frame:

$$\arg\min_{\theta} \mathcal{S}\Big(\Psi(\mathbf{I}_f) \circ \phi_2,\ \Psi(\mathbf{I}_m) \circ \phi_1\Big) \tag{1}$$

where $\mathcal{S}(\cdot,\cdot)$ represents the multi-scale MIND-SSC feature dissimilarity (sum of squared descriptor differences) [12]. The primary algorithmic challenges of this formulation entail: 1) estimating a parameter vector θ that resolves the geometric displacements across the sparse point cloud $\mathcal{P}$, and 2) interpolating these sparse matches onto the dense matrix grid Ω in a continuous, differentiable manner.

2.1 Transformation and Dense Interpolation Models

Deformable image registration frameworks frequently employ regular or irregular interpolation models that are analytically differentiable with respect to lower-dimensional spatial control fields. Prominent methodologies include grid-bound B-splines for implicit neural representations [17], log-Euclidean polyaffine frameworks [21], and differentiable Gaussian splatting for volumetric point cloud alignment [10]. Here, we unify log-Euclidean polyaffine modelling with point-cloud rasterisation, to yield an off-grid parameterisation that generalises classical grid-bound alternatives. This serves as a strong baseline for evaluating the impact of over-parameterisation, with or without neural fields, in sparse control-point-based registration. A critical limitation of traditional parametric registration is the reliance on a globally uniform smoothness assumption, which may either penalise local deformations too strongly or become unstable in uncertain regions. While spatially adaptive smoothing has been addressed via iterative Bayesian inference [19] or explicit multi-field optimisation grids [15], we introduce a dynamic, multi-scale blending framework embedded directly within a continuous interpolation architecture.

Let each sparse control point $\mathbf{p}_k \in \mathcal{P}$ carry a 12-DoF local transformation vector $\mathbf{M}_k \in \mathbb{R}^{12}$, which acts as an element of the log-Euclidean Lie algebra $\mathfrak{aff}(3)$ of affine transformations [1]. Concurrently, each control point is associated with an optimisable softmax mixture parameter $\pi_k \in \Delta^{M-1}$ across a discrete spectrum of M Gaussian scales $\mathcal{G} = \{\sigma_m\}_{m=1}^{M}$. To avoid computationally expensive, dense pairwise distance calculations when mapping from the unstructured point set $\mathcal{P}$ to the dense grid Ω, we leverage Differentiable Volumetric Rasterisation of Point Clouds (DiVRoC) [10]. We use differentiable forward splatting $\mathcal{S}$ to project sparse parameters from $\mathcal{P}$ onto the voxel grid Ω. The operation is applied to transformation parameters, scale probabilities π_k, and unit density weights $\mathbf{1}$, producing sparse voxel representations, where the values are distributed across the immediate eight trilinear neighbours in the discrete voxel grid.

$$\begin{aligned} &\mathbf{A}_{\text{sparse}}(\mathbf{x}) = \mathcal{S}(\{\mathbf{M}_k\}, \mathcal{P}), \quad \mathbf{\Pi}_{\text{sparse},m}(\mathbf{x}) = \mathcal{S}(\{\pi_{k,m}\}, \mathcal{P}), \\ &\mathbf{W}_{\text{sparse}}(\mathbf{x}) = \mathcal{S}(\{\mathbf{1}\}, \mathcal{P}) \end{aligned} \tag{2}$$

To enforce appropriate elastic regularity, these voxelised sparse representations are smoothed sequentially via a bank of separable 3D Gaussian filters G_{σ_m} assigned to each scale [15]. At each discrete scale σ_m, a normalised, scale-specific intermediate log-affine field $\mathbf{A}_m$ and its associated weight map ω_m are computed as follows via Nadaraya-Watson kernel regression:

$$\mathbf{A}_m(\mathbf{x}) = \frac{(G_{\sigma_m} * \mathbf{A}_{\text{sparse}})(\mathbf{x})}{(G_{\sigma_m} * \mathbf{W}_{\text{sparse}})(\mathbf{x}) + \epsilon}, \quad \omega_m(\mathbf{x}) = \frac{(G_{\sigma_m} * \mathbf{\Pi}_{\text{sparse},m})(\mathbf{x})}{(G_{\sigma_m} * \mathbf{W}_{\text{sparse}})(\mathbf{x}) + \epsilon} \tag{3}$$

Here $*$ denotes the spatial convolution operator and $\epsilon = 10^{-6}$ prevents numerical singularities.

Blending: Because varying Gaussian bandwidths diffuse spatial probabilities unevenly across coordinates, the dense weight maps $\omega_m(\mathbf{x})$ are explicitly normalised across the scale dimension:

$$\hat{\omega}_m(\mathbf{x}) = \frac{\omega_m(\mathbf{x})}{\sum_{j=1}^{M} \omega_j(\mathbf{x})} \tag{4}$$

The final, continuous log-affine matrix field $\mathbf{A}_{\text{dense}}(\mathbf{x})$ is constructed as a local convex combination across scales:

$$\mathbf{A}_{\text{dense}}(\mathbf{x}) = \sum_{m=1}^{M} \hat{\omega}_m(\mathbf{x}) \cdot \mathbf{A}_m(\mathbf{x}) \tag{5}$$

To construct symmetric midpoint mappings to a joint coordinate frame, we compute voxel-wise matrix exponentials via `torch.linalg.matrix_exp`:

$$\mathbf{T}_1(\mathbf{x}) = \exp\left(+\tfrac{1}{2}\mathbf{A}_{\text{dense}}(\mathbf{x})\right), \quad \mathbf{T}_2(\mathbf{x}) = \exp\left(-\tfrac{1}{2}\mathbf{A}_{\text{dense}}(\mathbf{x})\right) \tag{6}$$

These local transformation matrices map homogeneous grid coordinates $\mathbf{X}_h(\mathbf{x}) = [\mathbf{x}^T, 1]^T$ to yield deformations $\phi_1(\mathbf{x}) = \mathbf{T}_1(\mathbf{x})\mathbf{X}_h(\mathbf{x})$ and $\phi_2(\mathbf{x}) = \mathbf{T}_2(\mathbf{x})\mathbf{X}_h(\mathbf{x})$. While pointwise matrix exponentials guarantee local non-singularity ($\det(\mathbf{T}(\mathbf{x})) > 0$), spatial variation across voxels can introduce folding if unregularised and multi-scale Gaussian blending plays a key role in preserving spatial topology.

2.2 Over-Parameterisation via Implicit Neural Fields

While the multi-scale combination of log-Euclidean polyaffine matrices enables geometrically plausible deformations, optimising explicit parametric representations directly can make the registration highly sensitive to local minima, and the choice of smoothing parameters, e.g. $\mathcal{G}$, become very important. To avoid these constraints, SIREN-based neural fields have emerged as an implicit, higher-dimensional optimisation space for continuous deformable alignment [27]. While coordinate networks intrinsically model complex, continuous spatial variations, they often require carefully engineered pre-alignment initialisations [9] or coupled explicit smoothing functions (e.g., B-splines) [17] to reduce convergence issues and preserve topology. Departing from standard neural registration frameworks that evaluate coordinate MLPs globally across the dense matrix grid Ω, incurring high computational costs in 3D, we leverage the coordinate network as a continuous generative function operating exclusively over the unstructured control space $\mathcal{P}$. Formally, we define a Sinusoidal Representation Network (SIREN) $f_\theta : \mathbb{R}^3 \to \mathbb{R}^{12}$ parameterised by weights θ. The network maps each continuous control point coordinate $\mathbf{p}_k \in \mathcal{P}$ directly to its associated element in the log-Euclidean Lie algebra $\mathfrak{aff}(3)$:

$$\mathbf{M}_k = f_\theta(\mathbf{p}_k) \tag{7}$$

The hidden layers of f_θ leverage periodic activation functions of the form $\sin(\omega_0 \cdot \mathbf{z})$, where the hyperparameter ω_0 controls the input frequency scaling. This formulation introduces parameter-induced implicit regularisation during optimisation. Since the network weights θ are shared across all coordinates, the model

Table 1: Aggregate deformable registration performance comparing parameterisation paradigms across 50 inter-subject AMOS22 abdominal CT cases and 10 DIR-Lab COPD lung CT cases. Performance is reported via mean and standard deviation (± SD) of Dice Similarity Coefficient (DSC) and Target Registration Error (TRE), percentage fraction of negative Jacobian determinants ($|J| < 0$ in %), trainable parameter count (Params in k), and COPD execution runtime per case (Time in s).

	AMOS22 Abdominal CT			**COPD Lung CT**			
Method	**Mean Dice** (% ± SD)	**Neg. Jac.** (%)	**Params** (k)	**Mean TRE** (mm ± SD)	**Neg. Jac.** (%)	**Params** (k)	**Time** (s)
initial (unregistered)	20.8 ± 10.3	–	–	19.03 ± 5.90	–	–	–
deeds [11]	46.9 ± 8.6	–	–	8.20 ± 1.60	–	–	45.0
Baseline: 3-DoF Translation Field							
explicit Adam (1σ)	49.8 ± 10.8	0.0 ± 0.0	14.7	5.45 ± 3.15	0.0	4.0	16.7
explicit Adam (5σ)	50.5 ± 11.4	0.0 ± 0.0	14.7	2.92 ± 1.59	0.0	4.0	28.2
implicit INR (1σ)	52.0 ± 9.8	0.2 ± 0.4	199.2	2.76 ± 1.33	0.0	50.4	16.9
implicit INR (5σ)	52.7 ± 10.7	0.2 ± 0.4	398.9	2.73 ± 1.37	0.0	101.1	28.7
Proposed: 12-DoF Log-Affine Mixture							
explicit Adam (1σ)	52.3 ± 11.1	0.2 ± 0.2	59.0	5.69 ± 3.17	0.3	16.0	16.6
explicit L-BFGS (5σ)	48.3 ± 17.6	3.6 ± 10.8	59.0	12.63 ± 11.58	0.0	16.0	28.8
explicit Adam (5σ)	53.2 ± 11.1	0.2 ± 0.2	59.0	2.86 ± 1.57	0.0	16.0	28.3
implicit INR (5σ)	53.6 ± 10.0	0.9 ± 1.7	401.2	2.70 ± 1.39	0.0	102.3	28.7

naturally learns large-scale, low-frequency structure first, such as global scaling and overall affine transformations, before refining local, non-rigid deformations. This behaviour reduces the need for a separate rigid pre-alignment step.

3 Experiments and Results

We evaluate the influence of over-parameterisation on registration performance in two challenging settings: 1) inter-subject abdominal CT registration and 2) intra-subject large breathing motion estimation of lung CT. Our hypothesis is that, while directly optimising displacements using Adam may suffice for small, localised deformations, the introduction of either multi-scale log-Euclidean polyaffine modelling or a SIREN-based neural field improves convergence and registration quality under large initial misalignments.

3.1 Datasets and model configurations

We evaluate the proposed methods on the publicly available AMOS22 CT dataset [14], selecting nine organs aligned with previous benchmarks [4] and the Learn2Reg abdomenCT test set [13] (spleen, r. kidney, l. kidney, liver, stomach, aorta, vena cava, pancreas, and duodenum). For intra-subject lung registration, we use all 10 patient pairs of the DIR-Lab COPDgene dataset [3] resampled to isotropic 0.7 mm resolution, which provides inspiration and expiration scans incorporating 300 manual landmark pairs per case (with an initial mean TRE of $19.03 \pm 5.90\,\mathrm{mm}$). This dataset presents a substantial challenge compared to standard 4DCT data. We evaluate all 10 intra-patient pairs for COPDgene and 50 unique inter-subject pairings of AMOS22. We adapt the hyperparameter settings for MIND-SSC as recommended in [18], using two scales for the latter to ensure robustness to intensity variations. To focus on inner-lung structures, the loss functions are locally weighted using automatically extracted vessel masks for COPDgene (whereas the discrete deeds baseline operate globally over the lung bounding box). While domain-specific PDE registration frameworks such as pTVReg [25] achieve sub-millimetre accuracy tailored specifically for lung CT, our work evaluates a general off-grid over-parameterisation across multiple domains using an identical loss function with only moderate hyperparameter adaptation. To cope with the large initial misalignment in the AMOS22 dataset, we apply a symmetric log-Euclidean global affine pre-alignment by setting $|K| = 1$ and optimising a single 12-DoF matrix.

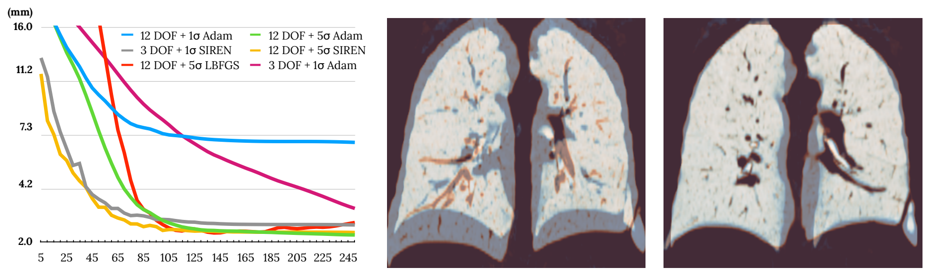
This pre-alignment step was unnecessary for the SIREN-based neural field approaches, as the global components within the coordinate network natively captures global affine motion during early iterations due to its spectral bias. For explicit models on COPDgene, we also used no global pre-alignment step, noticed however that it could be beneficial for severe COPD cases in future work. $|K|$ is set to 1,331 for COPDgene and 4,913 for AMOS22 to account for local deformation complexity. The multi-scale spatial regularisation uses five Gaussian kernels with $\sigma \in [6.0, 12.0]$ voxels or a single fixed $\sigma = 9$.

As a further baseline, we evaluate the discrete optimisation strategy **deeds** [11], which is designed for large-deformation alignment and outperformed [2] for abdominal CT registration. The SIREN architecture [20] was parameterised with either 4 layers and $d_{\text{hidden}} = 128$ or 256 for COPD and AMOS respectively, predicting 3 or 12 DoF per 3D coordinate. To ensure smooth deformation fields without high-frequency distortions around individual sparse points, we use a low initial frequency scaling ($\omega_0 = 2.5$) and project the bounded ($0.75\tanh(\cdot)$) elements into Gaussian splatting interpolation with either single fixed or multiple scales. When employing $M = 5$ probabilities π_k for locally dynamic adaptation of σ, we optimise a second SIREN with 5 output channels followed by a softmax. The parameters are trained with Adam for 150 iterations at $\eta = 3 \times 10^{-3}$ (using cosine annealing). We evaluate Adam and L-BFGS optimisers for the explicit over-parameterisation setup using 150 iterations and 30×5 steps ($\eta = 3 \times 10^{-3}$ and $\eta = 0.3$, respectively, without annealing).

All continuous models were implemented in PyTorch and benchmarked on a single NVIDIA A40 GPU (48 GB VRAM). Negative Jacobian determinant percentages ($|J| < 0$) are evaluated densely across the full volume bounding box.

3.2 Results and Discussion

The quantitative results in Table 1 demonstrate the effects of over-parameterisation from two clear viewpoints:



(a) Optimisation trajectory (b) Before registration (c) After registration

Fig. 2: Instance alignment diagnostics for COPD Case #6. (a) Deformable registration optimisation path (TRE vs. Iterations). Traditional 3-DoF translation fields converge slowly, while unconstrained 12-DoF single-scale models suffer from localised optimisation plateaus. Our proposed multi-scale explicit grid (12-DoF + 5σ) and the SIREN-based neural field both avoid these local minima, achieving rapid, robust error reduction. Visual anatomical alignment quality is verified by comparing the fixed-to-moving image overlay maps (b) before registration and (c) after non-rigid alignment.

1) explicit Adam instance optimisation substantially improves from extending the prediction to 12 DoF with multiple σ values, which reduces the TRE for COPD from 5.45 to 2.86 mm (a 16.17 mm reduction from initial error) and increases the Dice score for AMOS22 by 3.4 %pts. Single-scale explicit 12-DoF optimisation (1σ) degrades to 5.69 mm TRE and exhibits 0.3% negative Jacobians, showing that explicit over-parameterisation benefits from multi-scale smoothing (5σ) to avoid local minima and preserve spatial topology.

2) implicit neural optimisation utilises a larger parameter space but yields smaller incremental gains (0.2 mm and 1.6 %pts) from polyaffine locally adaptive prediction. Yet, it also incurs a larger negative Jacobian penalty. L-BFGS struggles to converge consistently, leading to highly irregular fields in five out of ten COPD cases [1]. Notably, the results of deeds [11] on lung CT are limited here due to the high-resolution nature of the isotropic volumes; adapting the data pre-processing or multi-level grid scales could potentially improve these results further. Overall, the resulting segmentation overlap of >53% for complex inter-subject abdominal registration from initially 21% is highly competitive and outperforms all unsupervised approaches explored in Learn2Reg [13].

The optimisation trajectories in Fig. 2 show that solver performance depends strongly on the registration parameterisation. With Adam, the baseline explicit model (3 DoF + 1σ Adam) converges slowly, while simply increasing flexibility (12 DoF + 1σ Adam) leads to poor convergence. Adding multi-scale smoothing (**12 DoF + 5σ Adam**) stabilises optimisation and achieves performance com-

[1] When comparing L-BFGS and SIREN with 12 DoF and 5σ only on cases # 1, 3, 6, 7, 9 both achieve similar TREs of 2.04 ± 0.60 mm and 2.02 ± 0.46 mm respectively, with 14.8 mm initial misalignment.

parable to the much larger SIREN-based neural field (**12 DoF + 5σ SIREN**) models (16.0k vs. 135.3k parameters on COPD). Although implicit neural representations converge faster initially due to their shared continuous parameterisation, careful explicit over-parameterisation largely closes this gap. On COPD, SIREN achieves a slight improvement over the explicit polyaffine model (2.70 mm vs. 2.86 mm); this small difference could stem from the fact that SIREN coordinate networks naturally captures global affine scaling and rotation across the entire volume during early iterations, whereas explicit control points were optimized without an initial global affine pre-alignment step. L-BFGS can also accelerate optimisation and escape some local minima, but in practice is less robust, frequently producing failed convergence or topologically implausible deformations.

4 Conclusion and Future Work

In this work, we systematically investigated the role of over-parameterisation in sparse, off-grid 3D medical image registration. By comparing explicit multi-scale log-Euclidean polyaffine frameworks against SIREN-based neural fields, we demonstrated that lifting the optimisation landscape, either via higher spatial degrees of freedom or via continuous architectural network weights, accelerates convergence and avoids poor local minima. Our findings indicate that while implicit neural fields provide an excellent global inductive bias for complex inter-subject alignment, explicit grid optimisation coupled with dynamic, multi-scale blending can achieve comparable precision for large-deformation estimation.

Building upon these findings, several compelling avenues remain for future exploration. First, we intend to explore image-conditioned dynamic keypoint or control point extraction methods, which would enable the sparse grid density to automatically adapt to local anatomical variations and topological changes. In addition, incorporating an explicit global affine pre-alignment step prior to local polyaffine optimization could further close the small performance gap to implicit neural fields on severe displacement cases. Second, integrating advanced medical or vision foundation features could provide vital semantic guidance, further bridging the performance gap in highly non-linear inter-subject registration tasks. Finally, we aim to analyse optimisation trajectories in greater detail to untangle the close coupling between second-order solvers, learning rate scheduling variants, parameter initialisation states, and the inherent landscape-smoothing effects of high-dimensional over-parameterisation.

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

References

1. Arsigny, V., Commowick, O., Pennec, X., Ayache, N.: A fast and log-euclidean polyaffine framework for locally linear deformable registration. Journal of Mathematical Imaging and Vision **33**(2), 222–238 (2009). https://doi.org/10.1007/s10851-008-0109-x

2. Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Medical image analysis **12**(1), 26–41 (2008)
3. Castillo, R., Castillo, E., Fuentes, D., Ahmad, M., Wood, A.M., Ludwig, M.S., Guerrero, T.: A reference dataset for deformable image registration spatial accuracy evaluation using the copdgene study archive. Physics in medicine and biology **58**(9), 2861–2877 (2013)
4. Gibson, E., Giganti, F., Hu, Y., Bonmati, E., Bandula, S., Gurusamy, K., Davidson, B., Pereira, S.P., Clarkson, M.J., Barratt, D.C.: Automatic multi-organ segmentation on abdominal ct with dense v-networks. IEEE transactions on medical imaging **37**(8), 1822–1834 (2018)
5. Hadramy, S.E., Cherkaoui, O., Cattin, P.C.: cIDIR: Conditioned implicit neural representation for regularized deformable image registration. In: Felsner, L., Küstner, T., Maier, A., Qin, C., Ahmadi, S.A., Kazi, A., Hu, X. (eds.) Reconstruction and Imaging Motion Estimation, and Graphs in Biomedical Image Analysis. pp. 87–96. Springer Nature Switzerland, Cham (2026)
6. Han, K., Sun, S., Yan, X., You, C., Tang, H., Naushad, J., Ma, H., Kong, D., Xie, X.: Diffeomorphic image registration with neural velocity field. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). pp. 1869–1879 (2023). https://doi.org/10.1109/WACV56688.2023.00191
7. Hansen, L., Heinrich, M.P.: Graphregnet: Deep graph regularisation networks on sparse keypoints for dense registration of 3d lung cts. IEEE Transactions on Medical Imaging **40**(9), 2246–2257 (2021)
8. van Harten, L.D., Stoker, J., Išgum, I.: Robust deformable image registration using cycle-consistent implicit representations. IEEE Transactions on Medical Imaging **43**(2), 784–793 (2024). https://doi.org/10.1109/TMI.2023.3321425
9. van Harten, L.D., Van Herten, R.L.M., Išgum, I.: REINDIR: Repeated embedding infusion for neural deformable image registration. In: Proceedings of the 7th International Conference on Medical Imaging with Deep Learning. Proceedings of Machine Learning Research, vol. 250, pp. 407–421. PMLR (2024)
10. Heinrich, M.P., Bigalke, A., Großbröhmer, C., Hansen, L.: Chasing clouds: differentiable volumetric rasterisation of point clouds as a highly efficient and accurate loss for large-scale deformable 3d registration. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 8026–8036 (2023)
11. 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)
12. Heinrich, M.P., Jenkinson, M.C., Bhushan, M., Matin, T., Gleeson, F.V., Brady, M., Schnabel, J.A.: Mind: Modality independent neighbourhood descriptor for multi-modal deformable registration. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). LNCS, vol. 7512, pp. 764–773. Springer (2012)
13. Hering, A., Hansen, L., Mok, T.C., Chung, A.C., Siebert, H., Häger, S., Lange, A., Kuckertz, S., Heldmann, S., Shao, 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 (2022)
14. Ji, Y., Bai, H., Ge, C., Yang, J., Zhu, Y., Zhang, R., Li, Z., Zhang, L., Ma, W., Wan, X., Luo, P.: AMOS: A large-scale abdominal multi-organ benchmark for versatile medical image segmentation. In: Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., Oh, A. (eds.) Advances in Neural Information Processing Systems. vol. 35, pp. 36722–36732. Curran Associates, Inc. (2022)

15. Li, J., Liu, X., Zhang, F., Cao, X., Buhmann, J.M., Zhang, Y., Li, X.: Gaussian primitives for deformable image registration. Physics and Imaging in Radiation Oncology p. 100821 (2025)
16. Mok, T.C.W., Li, Z., Xia, Y.s., Yao, J., Zhang, L., Zhou, J., Lu, L.: Deformable medical image registration under distribution shifts with neural instance optimization. In: Machine Learning in Medical Imaging – MLMI 2023. Lecture Notes in Computer Science, vol. 14348, pp. 126–136. Springer (2023). https://doi.org/10.1007/978-3-031-45673-2_13
17. Sideri-Lampretsa, V., McGinnis, J., Qiu, H., Paschali, M., Simson, W., Rueckert, D.: Sinr: Spline-enhanced implicit neural representation for multi-modal registration. In: Burgos, N., Petitjean, C., Vakalopoulou, M., Christodoulidis, S., Coupe, P., Delingette, H., Lartizien, C., Mateus, D. (eds.) Proceedings of The 7nd International Conference on Medical Imaging with Deep Learning. Proceedings of Machine Learning Research, vol. 250, pp. 1462–1474. PMLR (03–05 Jul 2024), https://proceedings.mlr.press/v250/sideri-lampretsa24a.html
18. Siebert, H., Großbröhmer, C., Hansen, L., Heinrich, M.P.: Convexadam: Self-configuring dual-optimization-based 3d multitask medical image registration. IEEE Transactions on Medical Imaging **44**(2), 738–748 (2024)
19. Simpson, I.J., Cardoso, M.J., Modat, M., Cash, D.M., Woolrich, M.W., Andersson, J.L., Schnabel, J.A., Ourselin, S., Initiative, A.D.N., et al.: Probabilistic non-linear registration with spatially adaptive regularisation. Medical image analysis **26**(1), 203–216 (2015)
20. Sitzmann, V., Martel, J., Bergman, A., Lindell, D., Wetzstein, G.: Implicit neural representations with periodic activation functions. Advances in neural information processing systems **33**, 7462–7473 (2020)
21. Taquet, M., Macq, B., Warfield, S.K.: Spatially adaptive log-euclidean polyaffine registration based on sparse matches. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 590–597 (2011)
22. Tian, L., Greer, H., Kwitt, R., Vialard, F.X., San Jose Estepar, R., Bouix, S., Rushmore, R., Niethammer, M.: uniGradICON: A foundation model for medical image registration. arXiv preprint arXiv:2403.05780 (2024). https://doi.org/10.48550/arXiv.2403.05780
23. Tian, L., Greer, H., San José Estépar, R., Sengupta, R., Niethammer, M.: NePhi: Neural deformation fields for approximately diffeomorphic medical image registration. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2024. Lecture Notes in Computer Science, vol. 15004, pp. 213–237. Springer (2024). https://doi.org/10.1007/978-3-031-73223-2_13
24. Velasquez, A., Jha, S., Alkhouri, I.R.: On the dataless training of neural networks. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 40, pp. 41013–41019 (2026)
25. 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 (2016)
26. de Vos, B.D., Berendsen, F.F., Viergever, M.A., Sokooti, H., Staring, M., Išgum, I.: A deep learning framework for unsupervised affine and deformable image registration. Medical Image Analysis **52**, 128–143 (2019). https://doi.org/10.1016/j.media.2018.11.010
27. Wolterink, J.M., Zwienenberg, J.C., Brune, C.: Implicit neural representations for deformable image registration. In: International Conference on Medical Imaging with Deep Learning. pp. 1349–1359. PMLR (2022)