Mirror and Map: Symmetric Latent Deformation Priors for Deformable Image Registration using Implicit Neural Representations

Bennet Kahrs[1[0009−0009−5609−9250]], Fenja Falta[1,3[0009−0003−3234−9725]], Jan Ehrhardt[1,2[0009−0004−6804−5587]], Heinz Handels[1,2[0000−0002−3499−4328]], and Timo Kepp[2[0000−0003−2024−2958]]

[1] Institute of Medical Informatics, University of Luebeck, Luebeck, Germany
[2] German Research Center for Artificial Intelligence, Luebeck, Germany
[3] Faculty of Electrical Engineering and Computer Science, Technical University of Applied Sciences Luebeck, Luebeck, Germany
bennet.kahrs@uni-luebeck.de

Abstract. Implicit neural representations (INRs) have recently emerged as a promising alternative to population-based CNNs and classical optimization-based methods for medical image registration, offering continuous, resolution-agnostic deformation fields. However, their instance-wise optimization can be prone to instability and implausible deformations. Generalizable INRs address this limitation by conditioning a shared network on subject-specific latent priors, yet their smoothness arises implicitly from the network rather than from an explicit constraint, offering no control over the regularity of the deformation. Moreover, existing generalizable INR frameworks typically require two independent latent priors to represent forward and backward deformations, although both describe the same underlying anatomical correspondence. We propose *Mirror and Map* (MnM), a symmetric, generalizable INR framework that encodes both deformation directions through a single latent prior and its negation. MnM combines a Lipschitz regularization scheme adapted to the modulated architecture with cycle-consistency regularization and Muon as an optimizer, jointly promoting smooth, plausible, and approximately diffeomorphic deformations. Compared with generalizable baselines, MnM improves registration accuracy while preserving deformation plausibility. We further show that the learned self-supervised deformation representation achieves classification performance comparable to image-based approaches. Code will be available at https://github.com/bkahrs/MirrorAndMap.

Keywords: Deformable Image Registration · Implicit Neural Representations · Self-Supervision · Representation Learning.

1 Introduction

Image registration describes the process of aligning two images by estimating a spatial transformation $\varphi_{R\to T}$ that maps a reference R onto a target T. This is an essential part in many medical applications, for example, to track anatomical changes across visits, capturing motion such as inhale and exhale phases of the lung, or aligning a subject to a common anatomical atlas [13,20]. Anatomical plausibility requires the transformation to be smooth and invertible, corresponding to a diffeomorphic mapping.

Data-driven approaches, such as convolutional neural networks (CNNs) automate this estimation by learning to predict deformations over a population, enabling fast inference [2]. However, these methods may reach their limits when generalizing to images with large or previously unknown deformations, and since they operate on a fixed grid, they are typically restricted to a single resolution [9,24]. Implicit neural representations (INRs), in contrast, parameterize the deformation as a continuous function that maps spatial coordinates to displacements [24,22]. This enables deformations at arbitrary resolution and yields transformations that are analytically differentiable [19]. However, such representations are commonly optimized per instance rather than learned over a population, making them sensitive to hyperparameter choices and optimization settings, which can lead to less plausible deformations [22].

To combine the benefits of the data-driven approaches and those of INRs, generalizable INRs [4] have been proposed for image registration to condition a shared network on a per-subject latent prior $\boldsymbol{z}$ [25,12]. In this setting, the shared network weights capture a structure common to all subjects, while each latent prior encodes subject-specific deformation information. Consequently, each transformation is represented by a separate latent prior. This requires learning two independent representations, $\boldsymbol{z}_{R\to T}$ and $\boldsymbol{z}_{T\to R}$, to model the mappings, $\varphi_{R\to T}$ and $\varphi_{T\to R}$. However, both deformations describe the same underlying anatomical transformation, except in opposite directions. This raises the natural question of how the two latent representations are related, and whether they can be reduced to a single non-redundant latent prior.

Meanwhile, several works have investigated regularization techniques and optimization strategies for instance-based INR across different domains, including image reconstruction and shape representations. While recent work studies Lipschitz regularization for instance-based INRs and reports promising results, including image registration, it does not investigate penalizing an upper bound of the Lipschitz constant [16,14] as a flexible way to directly constrain the Jacobian of the deformation [23]. Moreover, this line of work has not explored the use of such regularization in the context of generalizable INRs. Muon has recently been proposed as an optimizer for INRs and shown promising results, but its potential for deformable image registration and generalizable INRs remains unexplored [11,15].

Contribution We propose *Mirror and Map* (MnM), a symmetric generalizable INR framework that encodes opposite deformation directions through a single latent prior $\boldsymbol{z}$ and its negation $-\boldsymbol{z}$. The negation acts as the *Mirror*, inducing a symmetric latent parameterization, while the shared generalizable INR serves as the *Map*. Together with a Lipschitz regularization adapted to the generaliz-

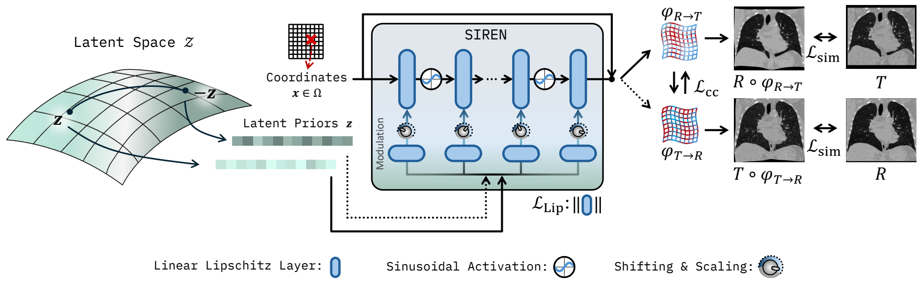

Fig. 1. Overview of the proposed method. The left side illustrates latent negation and the corresponding inputs to the shared network, while the right side shows the resulting image alignment.

able INR, Muon as an optimizer and a cycle-consistency term, MnM combines regularization and optimization techniques especially suited for image registration with INRs [14,11,15]. This promotes smooth and plausible deformations and yields approximately diffeomorphic registration.

2 Methods

Let $R_i : \Omega \to \mathbb{R},\ i = 1, \ldots, N$ denote a set of reference images and $T_i : \Omega \to \mathbb{R},\ i = 1, \ldots, N$ the corresponding set of target images, where N denotes the number of subjects and $\Omega = [-1, 1]^d$ represents the d-dimensional image domain. Symmetric and diffeomorphic image registration aims to capture anatomically plausible variations by estimating forward and backward deformation fields $\varphi_{R_i \to T_i}, \varphi_{T_i \to R_i} : \Omega \to \Omega$ such that

$$T_i \approx R_i \circ \varphi_{R_i \to T_i} \quad \text{and} \quad R_i \approx T_i \circ \varphi_{T_i \to R_i}. \tag{1}$$

Within this setting, *Mirror and Map* (MnM) builds on [12] and extends it with a symmetric latent constraint regularization, as illustrated in Fig. 1.

MnM's foundation is a single generalizable INR f_θ and trainable latent priors $\boldsymbol{z}_i \in \mathbb{R}^L$ as proposed in [4], where the weights θ and latent space $Z = (\boldsymbol{z}_i)_{i=1}^N \in \mathbb{R}^{N \times L}$ are optimized jointly over N subjects. Unlike the meta-learning scheme of [4], we treat the latent priors as trainable parameters and optimize them directly alongside the network weights at each update step. For each subject $i = 1, \ldots, N$, the network learns to align the reference images R_i to the target images T_i by modeling deformation fields

$$\varphi_i(\boldsymbol{x}) = f_\theta(\boldsymbol{x}, \boldsymbol{z}_i) + \boldsymbol{x},$$

where $\boldsymbol{x} \in \Omega$ denotes the spatial coordinates [12]. The latent priors $\boldsymbol{z}_i$ condition f_θ through feature-wise modulation applied to every hidden linear layer ℓ of f_θ [17,4]. For each layer, $\boldsymbol{z}_i$ is mapped by a separate linear layer to a scale vector

$\boldsymbol{\gamma}_i^{(\ell)}$ and a shift vector $\boldsymbol{\beta}_i^{(\ell)}$. For each hidden layer $\sigma(\omega(W^{(\ell)}\boldsymbol{h}^{(\ell)} + \boldsymbol{b}^{(\ell)}))$ with input $\boldsymbol{h}^{(\ell)}$, weight matrix $W^{(\ell)}$, and bias $\boldsymbol{b}^{(\ell)}$, this modulation yields

$$\sigma\big(\omega(\boldsymbol{\gamma}_i^{(\ell)} \odot W^{(\ell)}\boldsymbol{h}^{(\ell)} + \boldsymbol{b}^{(\ell)} + \boldsymbol{\beta}_i^{(\ell)})\big),$$

where $\odot$ denotes the element-wise multiplication, σ the sine activation function, and ω the SIREN scaling factor [19]. This way, the latent prior $\boldsymbol{z}_i$ encodes the deformation-specific information of subject i, while the shared parameters capture the general structure, which is common to all instances.

2.1 Symmetric Latent Constraints

Given that $\boldsymbol{z}_i$ encodes the forward deformation from R_i to T_i, we use its negated counterpart $-\boldsymbol{z}_i$ to represent the backward deformation. This construction induces a point symmetry in the latent space, where mirrored latent priors, $\boldsymbol{z}_i$ and $-\boldsymbol{z}_i$, encode opposite deformation directions of the same anatomical correspondence. Let $\varphi_{R_i \to T_i}$ denote the deformation that aligns the reference image R_i to the target image T_i, and $\varphi_{T_i \to R_i}$ the opposite direction. We impose a *symmetry constraint* by defining

$$\varphi_{R_i \to T_i}(\boldsymbol{x}) = f_\theta(\boldsymbol{x}, \boldsymbol{z}_i) + \boldsymbol{x}, \qquad \varphi_{T_i \to R_i}(\boldsymbol{x}) = f_\theta(\boldsymbol{x}, -\boldsymbol{z}_i) + \boldsymbol{x}.$$

To encourage invertible, approximately diffeomorphic deformations, we add a cycle-consistency regularization term. Inspired by GradICON [21], we combine an ICON [8] inverse-consistency term with a symmetric Jacobian determinant penalty of [22]. Since the INR is differentiable with respect to the spatial coordinates, the Jacobian can be evaluated analytically at arbitrary locations [24]. Writing the round-trip map as

$$\varphi_i^\circ(\boldsymbol{x}) = f_\theta\big(f_\theta(\boldsymbol{x}, \boldsymbol{z}_i) + \boldsymbol{x},\ -\boldsymbol{z}_i\big) + (f_\theta(\boldsymbol{x}, \boldsymbol{z}_i) + \boldsymbol{x}) = (\varphi_{T_i \to R_i} \circ \varphi_{R_i \to T_i})(\boldsymbol{x}),$$

the regularization loss is defined as

$$\mathcal{L}_{\mathrm{cc}}(\boldsymbol{z}_i) = \big\|\varphi_i^\circ(\boldsymbol{x}) - \boldsymbol{x}\big\|_1 + \lambda_{\mathrm{sJac}} \min\left\{ \left\| \frac{\big(\det \nabla\varphi_i^\circ(\boldsymbol{x}) - 1\big)^2}{\big|\det \nabla\varphi_i^\circ(\boldsymbol{x})\big|} \right\|_1,\ \tau \right\},$$

where $\det \nabla(\cdot)$ denotes the determinant of the Jacobian matrix, τ is a clipping threshold, λ_{sJac} the weighting parameter, and $\|\cdot\|_1$ the L1-norm. This formulation jointly penalizes the deviation from the identity transformation and encourages volume preservation during a full round-trip.

2.2 Lipschitz Regularization

To encourage both smooth and continuously differentiable deformations directly through the architecture, we build on a Lipschitz MLP [14] and adapt it to our modulated SIREN. With the latent $\boldsymbol{z}_i$ held fixed, f_θ is L-Lipschitz in the coordinate space if

$$\|f_\theta(\boldsymbol{x}, \boldsymbol{z}_i) - f_\theta(\boldsymbol{y}, \boldsymbol{z}_i)\|_\infty \leq L\,\|\boldsymbol{x} - \boldsymbol{y}\|_\infty.$$

Since the network is a MLP with Lipschitz-continuous sine activation, this can be achieved through specific normalization of the weight matrix $W^{(\ell)}$ using a trainable per-layer Lipschitz constant $c^{(\ell)}$ [14]. We adapt this to our network by normalizing the modulation-scaled weights $\boldsymbol{\gamma}_i^{(\ell)} \odot W^{(\ell)}$, such that the bound holds for every latent prior $\boldsymbol{z}_i$. We penalize the network-level Lipschitz bound during training via $\mathcal{L}_{\mathrm{Lip}} = \prod_\ell \omega c^{(\ell)}$, where ω denotes the activation scaling factor. This encourages the bound to remain small.

While this method was originally proposed to encourage Lipschitz continuity in the latent space, this regularization is particularly useful for deformation networks. By bounding the norm of the deformation's Jacobian, i.e., $\|\nabla f_\theta(\boldsymbol{x}, \boldsymbol{z}_i)\| \leq L$, it limits local stretching, resulting in smooth deformations [23].

2.3 Training

We take advantage of the symmetric construction by calculating the loss in both directions. Thus, we define a single directional loss for registering image A onto B with latent $\boldsymbol{z}$,

$$\mathcal{L}_{\mathrm{dir}}(A, B, \boldsymbol{z}) = \lambda_{\mathrm{sim}} \mathcal{L}_{\mathrm{sim}}\big(B,\ A \circ f_\theta(\cdot, \boldsymbol{z})\big) + \lambda_{\mathrm{cc}}\,\mathcal{L}_{\mathrm{cc}}(\boldsymbol{z}),$$

where λ_{sim} and λ_{cc} are weighting parameters and $\mathcal{L}_{\mathrm{sim}}$ is the normalized cross-correlation. We evaluate it for both directions by swapping the images and negating the latent priors:

$$\mathcal{L}_{\mathrm{total}} = \sum_{i=1}^{N} \mathcal{L}_{\mathrm{dir}}(R_i, T_i, \boldsymbol{z}_i) + \mathcal{L}_{\mathrm{dir}}(T_i, R_i, -\boldsymbol{z}_i) + \lambda_{\mathrm{Lip}}\,\mathcal{L}_{\mathrm{Lip}}.$$

where λ_{Lip} denotes a weighting parameter. During inference, the network weights θ are kept fixed, and only the latent prior $\boldsymbol{z}$ of the new subject is optimized. The inference loss reduces to $\mathcal{L}_{\mathrm{inf}} = \mathcal{L}_{\mathrm{dir}}(R, T, \boldsymbol{z}) + \mathcal{L}_{\mathrm{dir}}(T, R, -\boldsymbol{z})$.

3 Experiments and Results

We first analyze the structure of the learned latent embedding in a controlled toy setting, and then conduct two experiments on publicly available clinical datasets, followed by an ablation study. Beyond their registration capabilities, generalizable INRs can provide meaningful self-supervised representations for downstream analysis [7,5,1]. We therefore evaluate MnM with respect to both image registration performance and downstream pathology classification.

Preliminary Experiment We begin with a preliminary experiment to investigate the structure and learning behavior of the underlying latent priors. In an atlas registration setting, one volume containing two equally sized spheres is registered to a set of 300 volumes containing two spheres of varying size. The spheres are positioned at the same location across all volumes, while their radii are randomly sampled. Fig. 2 illustrates the experimental setup for MnM and compares the results with the non-symmetric baseline using the first two principal components. MnM produces a cleaner and more meaningful latent embedding, particularly for the latent priors that encode changes in both spheres, as indicated by green markers (LR). The latent priors corresponding to changes in the left or right sphere are structured along two orthogonal axes, while the priors corresponding to changes in both spheres are positioned consistently with these axes.

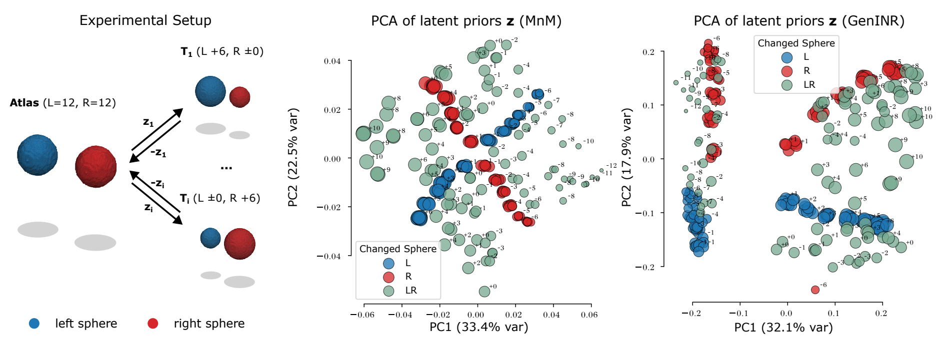

Fig. 2. Insights into the latent representations, illustrated using a toy example with volumes containing two spheres. Left: experimental setup showing two examples with differently sized spheres, where the radius of each sphere is changed by up to ± 6 voxels relative to the atlas. Middle: PCA of the latent embedding produced by MnM. Right: PCA of the latent embedding produced by the non-symmetric baseline. Latent priors are shown as markers and color-coded according to the sphere undergoing a volume change: left sphere (L), right sphere (R), or both spheres (LR). The absolute radius change is indicated by the marker size and label. Overlapping labels were removed for better readability.

Datasets We evaluate our method on two publicly available datasets. To assess registration accuracy under large deformations, we use the Lung250M-4B [6] dataset of paired in- and expiratory lung CT data. We hold out all cases with landmarks as the test set and use the remainder for training, yielding a 97/27 train-test split, and report the Dice score of the vessel segmentation and the target registration error (TRE).

To assess registration accuracy and, in particular, self-supervised representation quality, we use the ADNI1 dataset, containing pathology labels related to morphological changes [3]. We selected the first available time point per-patient and included only patients labeled with Alzheimer's disease (AD) or cognitively normal (CN), giving 243 training, 82 validation, and 82 test patients. All patients were skull-stripped and pre-aligned to the SRI24 atlas [18] using a rigid transformation. We report SSIM and the Dice score over six annotated structures between the SRI24 atlas and the target subjects for registration accuracy and use the AD/CN labels for the downstream classification task.

Implementation Details We parameterize f_θ as a SIREN [19] with three hidden layers of size 256, scaling parameter $\omega = 30$, and a latent prior of size $L = 16{,}000$ to better capture the variability of the lung vessels. We initialize $\theta \sim \mathcal{N}(0, 0.001)$ and $\boldsymbol{z_i} \sim \mathcal{N}(0, 0.01)$. The model is trained for 500 epochs using the Muon optimizer [11,15] with a cosine-annealed learning rate of 10^{-3} and a batch of 20,000 sampled coordinates. We use a weighted loss with $\lambda_{\text{sim}} = 10$, $\lambda_{\text{cc}} = 6$, $\lambda_{\text{sJac}} = 1$, and $\lambda_{\text{Lip}} = 10^{-4}$. For Lung250M-4B, coordinates are sampled uniformly within the lung mask. For ADNI, the coordinates are drawn from a normal distribution over the full image domain. At inference, the network weights are kept frozen, and only the latent prior is optimized for 300 iterations with a learning rate of 10^{-2} and a batch size of 100,000 sampled coordinates.

Registration Performance To assess registration performance, we compare MnM with the baseline generalizable INR (GenINR) [12], a modulated SIREN using non-symmetric latent priors. We further compare against two instance-based approaches, IDIR [24] and CycleIDIR [22]. Quantitative results are reported in Tab. 1 and Fig. 3.

On Lung250M-4B, MnM significantly improves registration accuracy compared with the baseline GenINR, reducing the TRE from 4.87 mm to 3.36 mm and increasing the vessel Dice score from 46.41% to 57.84%. With this TRE, MnM also achieves better alignment than the reported results of the learning-based VoxelMorph++ of 4.3 mm [10,6]. The instance-optimized methods IDIR and CycleIDIR achieve a lower TRE (2.69 and 2.65 mm) and higher Dice (68.43% and 71.64%), reflecting their highly subject-specific alignment of the lung vessels. However, this accuracy comes at the cost of increased folding and reduced de-

Table 1. Quantitative comparison of registration methods on both datasets. Metrics are stated as $\text{mean}_{\pm\text{standard deviation}}$. The best scores are highlighted in bold.

	Lung250M-4B			ADNI		
Method	TRE [mm]	DICE [%]	Neg. Jac. [%]	DICE [%]	SSIM [%]	Neg. Jac. [%]
IDIR	$2.69_{\pm 2.18}$	$68.43_{\pm 11.18}$	$0.95_{\pm 1.21}$	$\mathbf{90.34_{\pm 1.30}}$	$93.20_{\pm 0.77}$	$0.18_{\pm 0.52}$
CycleIDIR	$\mathbf{2.65_{\pm 2.06}}$	$\mathbf{71.64_{\pm 9.05}}$	$1.48_{\pm 1.87}$	$\mathbf{90.34_{\pm 1.30}}$	$\mathbf{94.14_{\pm 0.68}}$	$0.47_{\pm 0.56}$
GenINR	$4.87_{\pm 2.22}$	$46.41_{\pm 9.77}$	$\mathbf{0.00_{\pm 0.00}}$	$88.47_{\pm 1.35}$	$91.75_{\pm 1.04}$	$0.03_{\pm 0.08}$
MnM	$3.36_{\pm 0.84}$	$57.84_{\pm 7.21}$	$\mathbf{0.00_{\pm 0.00}}$	$88.54_{\pm 1.47}$	$91.93_{\pm 0.98}$	$\mathbf{0.00_{\pm 0.01}}$

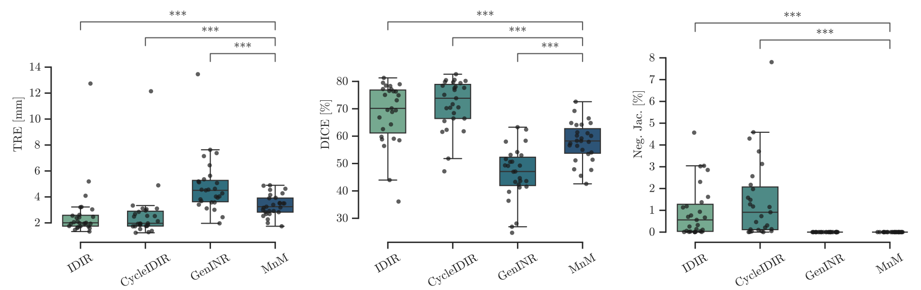

Fig. 3. Registration performance on the Lung250M-4B dataset. Significance was assessed with the Wilcoxon signed-rank test and Holm–Bonferroni correction (*** indicates $p < 0.001$).

formation plausibility, with negative Jacobian determinants of 0.95% and 1.48% compared with 0.00% for MnM.

Beyond the aggregated metrics, Fig. 3 shows that MnM yields more robust and more plausible deformations on Lung250M-4B. It produces no folding within the lungs and exhibits a markedly tighter TRE distribution with fewer outliers, suggesting greater resilience to domain shift and large deformations. This is reflected in the lowest TRE and Dice standard deviations among all methods (± 0.84 mm and $\pm 7.21\%$).

On ADNI, MnM performs on par with GenINR across the six labeled structures, with no significant difference in Dice (88.54% vs. 88.47%). MnM again exhibits less folding than all comparison methods on ADNI (0.00% vs. up to 0.47%), underscoring the robustness of the deformation.

Expressiveness of Latent Priors Complementing the registration evaluation, we investigate the representational capacity of the learned latent embeddings through downstream classification [4,7]. We compare the classification based on the self-supervised latent embeddings learned by MnM and GenINR with an image-based 3D-ResNet18 for Alzheimer's disease classification.

MnM and GenINR are first trained in a self-supervised manner for atlas-registration using the SRI24 atlas [18]. We then train Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM) classifiers on the corresponding training-set latent embeddings. Before classification, we apply principal component analysis (PCA) for dimensionality reduction, fitted on the training set, retaining 84 components explaining approx. 80% of the variance.

Tab. 2 reports classification performance on instance-optimized latent priors from the retained test set. For both classifiers, MnM improves performance across all metrics and achieves the highest accuracy of 0.756 with SVM. Notably, MnM's self-supervised deformation representation performs on par with an image-based 3D-ResNet18.

Table 2. AD/CN classification on the ADNI test set. The best results are highlighted.

	LDA		SVM		3D-ResNet18
Metric	MnM	GenINR	MnM	GenINR	Image
AUC	0.813	0.730	0.804	0.736	**0.814**
Accuracy	0.720	0.659	**0.756**	0.683	0.732
F1	0.710	0.658	**0.752**	0.683	0.732

Ablation Study We assess the contribution of MnM's components through an ablation study, reported in Tab. 3. Starting from the non-symmetric baseline GenINR, we successively add components: the symmetric latent negation, Muon [11,15], Lipschitz regularization and cycle-consistency. Registration performance improves steadily with each component, and the final method achieves a sub-voxel cycle error of 0.86 mm at 1 mm isotropic spacing. The Muon optimizer further improves training stability and registration performance for both GenINR and MnM, reaching mean TRE values of 3.91 mm and 3.85 mm. Notably, Muon yields a particularly strong performance gain for the non-symmetric, less-regularized GenINR, highlighting the potential of orthogonal optimizers [15].

We further isolate the effect of Lipschitz weight normalization [14]. Replacing IDIR's gradient-based regularization with Lipschitz weight normalization reduces negative Jacobian determinants from approximately 0.95% to 0.01% while maintaining a TRE of 2.67 mm. It also reduces runtime by 2.36 minutes (−82% reduction) per subject, suggesting improved deformation plausibility and efficiency, in line with recent findings on Lipschitz regularization for shape-representation INRs [16].

Table 3. Ablation of Lipschitz regularization (Lip.), cycle-consistency (ccLoss) and Muon optimizer on Lung250M-4B. Cycle-consistency error (ccError) denotes the mean displacement error after one round-trip. The best results are highlighted in bold.

Method	Muon	Lip.	ccLoss	TRE [mm]	Neg. Jac. [%]	ccError [mm]
GenINR (Map)	–	–	–	$4.87_{\pm 2.22}$	$\mathbf{0.00}_{\pm \mathbf{0.00}}$	–
	✓	–	–	$3.91_{\pm 1.39}$	$0.00_{\pm 0.01}$	–
Mirror & Map	–	–	–	$4.35_{\pm 1.50}$	$0.00_{\pm 0.02}$	$2.44_{\pm 1.27}$
	✓	–	–	$3.85_{\pm 1.17}$	$\mathbf{0.00}_{\pm \mathbf{0.00}}$	$2.19_{\pm 1.41}$
	✓	✓	–	$3.44_{\pm 0.86}$	$\mathbf{0.00}_{\pm \mathbf{0.00}}$	$1.97_{\pm 1.11}$
	✓	✓	✓	$\mathbf{3.36}_{\pm \mathbf{0.84}}$	$\mathbf{0.00}_{\pm \mathbf{0.00}}$	$\mathbf{0.86}_{\pm \mathbf{0.43}}$

4 Conclusion

This work introduced *Mirror and Map* (MnM), a Lipschitz-continuous generalizable INR for approximately diffeomorphic image registration based on a symmetric latent constraint. MnM exploits the redundancy between forward and inverse registration by encoding both directions through a point-symmetric latent representation. Compared with a single-directional GenINR, MnM improves registration performance while providing a more efficient and meaningful latent representation that performs on par with an image-based classifier. Moreover, MnM achieves greater deformation plausibility and robustness than both generalizable and instance-based INR baselines. In addition, we introduced Lipschitz regularization of the deformation field, improving deformation plausibility and robustness. Developing regularization strategies that jointly control the spatial deformation field and the latent representation remains an important direction for future work.

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

References

1. Aulakh, A., Forkert, N.D., Wilms, M.: Semi-disentangled spatiotemporal implicit neural representations of longitudinal neuroimaging data for trajectory classification. In: Hou, B., Mathai, T.S. (eds.) Learning with Longitudinal Medical Images and Data, vol. 16184, pp. 34–45. Springer Nature Switzerland (2026). https://doi.org/10.1007/978-3-032-16128-4_4
2. Chen, J., Liu, Y., Wei, S., Bian, Z., Subramanian, S., Carass, A., Prince, J.L., Du, Y.: A survey on deep learning in medical image registration: New technologies, uncertainty, evaluation metrics, and beyond. Medical Image Analysis **100**, 103385 (2025). https://doi.org/10.1016/j.media.2024.103385
3. Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M.O., Chupin, M., Benali, H., Colliot, O.: Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database. NeuroImage **56**(2), 766–781 (2011). https://doi.org/10.1016/j.neuroimage.2010.06.013
4. Dupont, E., Kim, H., Eslami, S.M.A., Rezende, D.J., Rosenbaum, D.: From data to functa: Your data point is a function and you can treat it like one. In: Proceedings of the 39th International Conference on Machine Learning. pp. 5694–5725. PMLR (2022)
5. Engelson, S., Kahrs, B., Kepp, T., Andresen, J., Handels, H., Ehrhardt, J.: Revealing and reducing morphological biases using implicit neural representations for medical image registration. In: Proceedings of The 9th International Conference on Medical Imaging with Deep Learning. pp. 4026–4041. PMLR (2026)
6. Falta, F., Großbröhmer, C., Hering, A., Bigalke, A., Heinrich, M.P.: Lung250M-4B: A combined 3D dataset for CT- and point cloud-based intra-patient lung registration. In: Proceedings of the 37th International Conference on Neural Information Processing Systems. pp. 54819–54832. NIPS '23, Curran Associates Inc. (2023)

7. Friedrich, P., Bieder, F., McGinnis, J., Wolleb, J., Rueckert, D., Cattin, P.C.: MedFuncta: A unified framework for learning efficient medical neural fields. In: Proceedings of The 9th International Conference on Medical Imaging with Deep Learning. pp. 56–87. PMLR (2026)
8. Greer, H., Kwitt, R., Vialard, F.X., Niethammer, M.: ICON: Learning regular maps through inverse consistency. In: 2021 IEEE/CVF International Conference on Computer Vision (ICCV). pp. 3376–3385. IEEE (2021). https://doi.org/10.1109/ICCV48922.2021.00338
9. Hansen, L., Heinrich, M.P.: Revisiting iterative highly efficient optimisation schemes in medical image registration. In: De Bruijne, M., Cattin, P.C., Cotin, S., Padoy, N., Speidel, S., Zheng, Y., Essert, C. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2021, vol. 12904, pp. 203–212. Springer International Publishing (2021). https://doi.org/10.1007/978-3-030-87202-1_20
10. Heinrich, M.P., Hansen, L.: Voxelmorph++: Going Beyond the Cranial Vault with Keypoint Supervision and Multi-channel Instance Optimisation. In: Hering, A., Schnabel, J., Zhang, M., Ferrante, E., Heinrich, M., Rueckert, D. (eds.) Biomedical Image Registration, vol. 13386, pp. 85–95. Springer International Publishing (2022). https://doi.org/10.1007/978-3-031-11203-4_10
11. Jordan, K., Jin, Y., Boza, V., You, J., Cesista, F., Newhouse, L., Bernstein, J.: Muon: An optimizer for hidden layers in neural networks (2024), https://kellerjordan.github.io/posts/muon/
12. Kahrs, B., Andresen, J., Falta, F., Santarossa, M., Handels, H., Kepp, T.: Don't Mind the Gaps: Implicit Neural Representations for Resolution-Agnostic Retinal OCT Analysis. Machine Learning for Biomedical Imaging **2026**(MELBA–BVM 2025 Special Issue), 59–77 (2026). https://doi.org/10.59275/j.melba.2026-38ba
13. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B., Sánchez, C.I.: A survey on deep learning in medical image analysis. Medical Image Analysis **42**, 60–88 (2017). https://doi.org/10.1016/j.media.2017.07.005
14. Liu, H.T.D., Williams, F., Jacobson, A., Fidler, S., Litany, O.: Learning Smooth Neural Functions via Lipschitz Regularization. In: Special Interest Group on Computer Graphics and Interactive Techniques Conference Proceedings. pp. 1–13. ACM (2022). https://doi.org/10.1145/3528233.3530713
15. McGinnis, J., Hölzl, F.A., Shit, S., Bieder, F., Friedrich, P., Mühlau, M., Menze, B., Rueckert, D., Wiestler, B.: Optimizing rank for high-fidelity implicit neural representations. In: Forty-Third International Conference on Machine Learning (2026)
16. McGinnis, J., Shit, S., Hölzl, F.A., Friedrich, P., Büschl, P., Sideri-Lampretsa, V., Mühlau, M., Cattin, P.C., Menze, B., Rueckert, D., Wiestler, B.: Beyond Uniformity: Regularizing Implicit Neural Representations through a Lipschitz Lens. In: The Fourteenth International Conference on Learning Representations (2026)
17. Perez, E., Strub, F., De Vries, H., Dumoulin, V., Courville, A.: FiLM: Visual reasoning with a general conditioning layer. In: AAAI Conf. Artif. Intell., AAAI. pp. 3942–3951. AAAI press (2018)
18. Rohlfing, T., Zahr, N.M., Sullivan, E.V., Pfefferbaum, A.: The SRI24 multichannel atlas of normal adult human brain structure. Human Brain Mapping **31**(5), 798–819 (2010). https://doi.org/10.1002/hbm.20906
19. Sitzmann, V., Martel, J.N.P., Bergman, A.W., Lindell, D.B., Wetzstein, G.: Implicit neural representations with periodic activation functions. In: Proceedings of the 34th International Conference on Neural Information Processing Systems. pp. 7462–7473. NIPS '20, Curran Associates Inc. (2020)

20. Sotiras, A., Davatzikos, C., Paragios, N.: Deformable medical image registration: A survey. IEEE Transactions on Medical Imaging **32**(7), 1153–1190 (2013). https://doi.org/10.1109/TMI.2013.2265603
21. Tian, L., Greer, H., Vialard, F.X., Kwitt, R., Estépar, R.S.J., Rushmore, R.J., Makris, N., Bouix, S., Niethammer, M.: GradICON: Approximate Diffeomorphisms via Gradient Inverse Consistency. In: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 18084–18094. IEEE (2023). https://doi.org/10.1109/CVPR52729.2023.01734
22. 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
23. Virmaux, A., Scaman, K.: Lipschitz regularity of deep neural networks: Analysis and efficient estimation. In: Advances in Neural Information Processing Systems. vol. 31. Curran Associates, Inc. (2018)
24. 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. pp. 1349–1359. PMLR (2022)
25. Zimmer, V.A., Hammernik, K., Sideri-Lampretsa, V., Huang, W., Reithmeir, A., Rueckert, D., Schnabel, J.A.: Towards generalised neural implicit representations for image registration. In: Mukhopadhyay, A., Oksuz, I., Engelhardt, S., Zhu, D., Yuan, Y. (eds.) Deep Generative Models, vol. 14533, pp. 45–55. Springer Nature Switzerland (2024). https://doi.org/10.1007/978-3-031-53767-7_5