NePhi: Neural Deformation Fields for Approximately Diffeomorphic Medical Image Registration

Lin Tian[1], Hastings Greer[1], Raúl San José Estépar[2], Roni Sengupta[1], and Marc Niethammer[1]

[1] UNC Chapel Hill
[2] Harvard Medical School

Abstract. This work proposes NePhi, a generalizable neural deformation model which results in approximately diffeomorphic transformations. In contrast to the predominant voxel-based transformation fields used in learning-based registration approaches, NePhi represents deformations functionally, leading to great flexibility within the design space of memory consumption during training and inference, inference time, registration accuracy, as well as transformation regularity. Specifically, NePhi 1) requires less memory compared to voxel-based learning approaches, 2) improves inference speed by predicting latent codes, compared to current existing neural deformation based registration approaches that *only* rely on optimization, 3) improves accuracy via instance optimization, and 4) shows excellent deformation regularity which is highly desirable for medical image registration. We demonstrate the performance of NePhi on a 2D synthetic dataset as well as for real 3D medical image datasets (e.g., lungs and brains). Our results show that NePhi can match the accuracy of voxel-based representations in a single-resolution registration setting. For multi-resolution registration, our method matches the accuracy of current SOTA learning-based registration approaches with instance optimization while reducing memory requirements by a factor of five. Our code is available at https://github.com/uncbiag/NePhi.

1 Introduction

Image registration is a key task in medical image analysis [11, 44]. The goal of image registration is to establish spatial correspondences between image pairs. The ideal registration algorithm should require minimal memory, should be fast and accurate, and should provide plausible transformations. However, while many registration algorithms have been developed over the last several decades [3, 7, 20, 30, 50, 55, 60], they typically only provide a subset of these desired qualities. As a consequence, different registration tasks will require different registration algorithms. The goal of this work is to provide NePhi, a neural registration algorithm that allows for significant freedom to adjust registration algorithm properties within the design space of memory consumption, speed, accuracy, and deformation regularity as shown in Fig. 1.

Memory consumption: Classical optimization-based registration algorithms [1, 19–21, 42, 46, 55] that optimize directly over the parameters of a transformation model (e.g.,

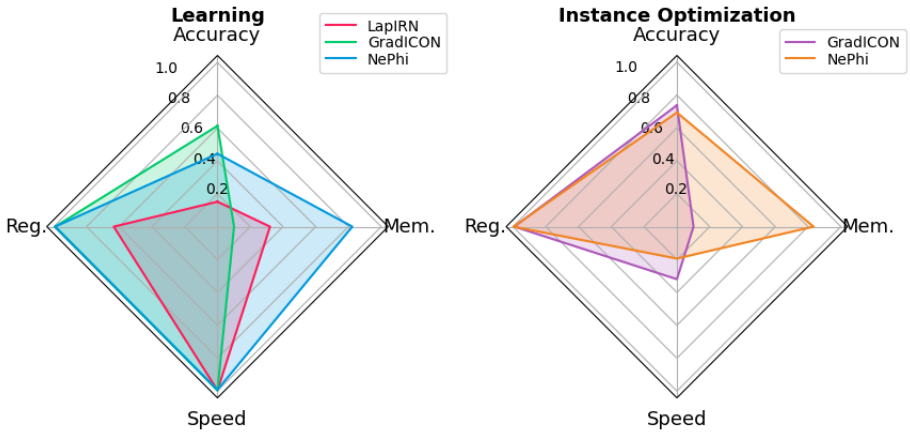

Fig. 1: Comparison between `NePhi` and competing methods based on four design considerations: registration accuracy, regularity of the predicted deformation field, peak memory consumption during training, and inference run-time. `NePhi` performs slightly worse for pure prediction but on-par with GradICON when combined with instance optimization. In both cases, `NePhi` shows better deformation regularity and memory consumption compared to other methods. Values are normalized to [0,1] for visualization. Refer to Supp. B.1 for details.

a 3D displacement field) are highly memory efficient. However, current registration research is focused on registration approaches based on deep learning due to achievable improvements in registration accuracy and speed [4, 7, 12, 16, 29, 30, 41, 50, 60]. While directly predicting a voxel-represented displacement field with a deep network is fast, memory consumption scales with the number of voxels and can therefore quickly become prohibitive for large 3D image volumes. This then necessitates the use of downsampled images and small batch sizes during training even on modern high-memory GPUs [7, 16, 30, 50]. Further, to achieve best performance these approaches perform instance-optimization [50] at test time where the neural network parameters are optimized for a given image pair. Hence, memory requirements are high during training *and* testing making these approaches unsuitable for low-memory GPUs, high-resolution images, or high-dimensional vector-valued images. Hence, exploring alternative transformation parameterizations is desirable to reduce memory consumption; this is expected to become increasingly important, e.g., for very high resolution volumetric microscopy images. *Our `NePhi` approach allows for low memory consumption during training and testing as it is based on an implicit neural representation parameterized by a multi-layer perceptron (MLP). Further, it is expected that such a representation can be parameter-efficient by exploiting the expectation that deformation spaces for real data are low-dimensional. In contrast, direct parameterization of a displacement field is very high dimensional; we would for example have 3 million degrees of freedom for a* 100^3 *image in 3D.*

Inference speed and accuracy: Voxel-based deep network approaches are fast but only achieve high accuracies with memory-intensive instance optimization [50]. Im-

plicit neural representations of deformations [5, 17, 18, 45, 53, 57, 58, 63], which are parameterized by MLPs, can achieve high accuracy, but are slow as they optimize MLPs from scratch for each image pair. *Our approach allows balancing inference time and accuracy because it predicts global and latent codes on which the MLP is conditioned. Hence a registration solution can be obtained by pure prediction or, if additional inference time is acceptable, can be refined for additional accuracy based on instance optimization.*

Regularity: In medical image registration, transformation regularity is often important and achieved by adding an appropriate regularizer, for example, based on concepts of diffusion [49], curvature [22], or models from continuum mechanics [5, 9, 10, 54]. *Our approach provides highly regular transformations by using gradient inverse consistency regularization, a regularizer which has shown to be highly effective for voxel-based deep registration models [50]. We further show that using this regularizer only on a very sparse set of locations is sufficient to regularize* `NePhi` *thereby further improving memory efficiency.*

The contributions of our work are as follows:

1) We propose a generalizable neural deformation field approach, `NePhi`, which is based on an implicit neural representation via an MLP. In contrast to existing neural deformation models for optimization-based image registration [17, 18, 45, 57, 58, 63], `NePhi` conditions the MLP on global and local codes thereby allowing to *predict* transformations.
2) We demonstrate that `NePhi` can provide high accuracy registration results at significantly lower memory cost than voxel-based deep registration networks.
3) While our base `NePhi` model directly predicts a deformation field, we also provide a hybrid multi-step approach for improved performance. The first step of this multi-step approach uses a low-resolution voxel-based model, which can be computed efficiently thereby retaining the overall efficiency of our approach.
4) By integrating a gradient inverse consistency (GradICON) loss into `NePhi` we obtain approximately diffeomorphic transformations (i.e., transformations that are smooth, have a smooth inverse, and are bijective) even when only evaluating this loss at a sparse set of spatial locations. Going beyond the existing GradICON approach we propose approximately diffeomorphic neural deformation fields (NDF) which represent the forward and backward deformations with *one* set of latent codes that can be predicted by a CNN in one forward pass.
5) We demonstrate the performance of `NePhi` on synthetic data, a real 3D lung registration between inhale and exhale computed tomography (CT) images, and a real brain MRI dataset. We evaluate our method across our four critical design considerations (memory consumption, accuracy, regularity, and training and inference time). This evaluation shows that our approach i) is very memory efficient, ii) can achieve excellent registration performance when combined with instance optimization or alternatively quickly predict registrations at reduced accuracy; and iii) results in highly regular transformations.

2 Related work

2.1 Medical Image Registration

Image registration is classically formulated as an optimization problem [1,19,21,42,46] with respect to a pair of images. The goal is to estimate optimal transformation parameters that minimize a compromise between transformation regularity and the dissimilarity between the transformed moving image with respect to the fixed image. As generally no closed-form solutions to these optimization problems exist they are solved using iterative numerical optimization (typically using a gradient descent scheme), resulting in significant computation time. To improve capture range multi-resolution approaches have been proposed [2, 27]. To reduce registration run-time, learning-based methods [4, 12, 16, 29, 30, 41, 50, 60] have been developed which replace numerical optimization by regression via a deep neural network at inference time. The pioneering works [12, 60] demonstrated how convolutional neural networks (CNN) can be used for learning-based registration. Subsequent developments added multi-step and multi-resolution capabilities [16, 29, 30, 50] to improve registration accuracy; these approaches achieve competitive accuracies to optimization-based multi-resolution registration methods but with significantly less run-time. Instance-optimization, where the registration networks are finetuned for a given image pair, can further increase registration performance [50]. While learning-based multi-resolution registration methods provide state-of-the-art performance and nearly instantaneous inference, they consume significant memory during training and instance optimization (IO), because the used CNNs predict dense deformation fields and feature maps which becomes especially costly in 3D. Registrations at high spatial resolutions for 3D image volumes are therefore not easily possible. Further, these approaches are relatively slow when using IO as large neural networks are finetuned. *Our motivation is therefore to develop* `NePhi`*, a neural deformation model which has a better memory profile and allows balancing between runtime, memory consumption, and registration accuracy.*

2.2 Neural Deformation Models

Functional representations of deformation fields (parameterized via MLPs) for natural images in dynamic scenes [14,24,25,25,33,34,36,51], for dynamic objects [13,23,31, 47, 56], and for animatable humans [8, 15, 26, 35, 40, 59, 61, 62] have attracted significant recent interest. These approaches focus on reconstructing an underlying 3D scene, object, or human while accounting for existing motion between acquired 2D images. However, only limited work exists on using such functional representations for medical image registration [17, 18, 45, 53, 57, 58, 63]. Existing approaches either use an MLP-parameterized function to directly represent a displacement vector field (DVF) [57, 63] or to represent velocity or momentum fields [17, 45, 58] to capture large deformations. In velocity or momentum field approaches a transformation is obtained by numerical integration; this indirect parameterization ensures diffeomorphic transformations. but is computationally costly because of the required numerical integrations. *Most importantly, all existing registration work based on neural deformation models only focuses*

on optimization-based *registration, which is slow compared to learning-based registration which predicts transformations at test time. In contrast, our proposed* NePhi *approach is suitable for optimization-based* and *learning-based registration.*

3 Background

Given two images $I^A : \Omega^A \to \mathbb{R}$ and $I^B : \Omega^B \to \mathbb{R}$, where $\Omega^{A,B} \in \mathbb{R}^d$ indicates the domain of the images, our goal is to find a mapping $\varphi^{AB} : \mathbb{R}^d \to \mathbb{R}^d$ such that $I^A \circ \varphi^{AB} \approx I^B$[3]. This continuous transformation map, φ^{AB}, can be parameterized as a function with parameters τ, denoted as $\varphi^{AB}(x;\tau)$, where the parameters depend on the function class to be parameterized. E.g., these could be parameters for an affine transform $\varphi^{AB}(x;\tau) = Tx + b$ where $\tau = \{T, b\}$, a discretized vector field, τ, on a voxel grid $\varphi^{AB}(x;\tau) = x+$ interpolate(τ at x), or the condition latent code z to an MLP $\varphi^{AB}(x;\tau) = \varphi_\theta(x;z)$ with $\tau = z$ and φ_θ being the MLP.

Learning-based registration. Here, one trains a neural network (NN) $f_\theta(I_i^A, I_i^B) = \tau$ that outputs the parameters of a chosen transformation model φ. We denote the predicted transformation field as $\varphi(x;\tau)$. The NN is trained over a set of image pairs $I = \{(I_i^A, I_i^B)\}_{i=1}^N$ by solving

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}_{\text{sim}}\left(I_i^A \circ \varphi(x;\tau), I_i^B\right) + \lambda \mathcal{L}_{\text{reg}}(\varphi(x;\tau)), \tau = f_\theta(I_i^A, I_i^B), \quad (1)$$

where $\mathcal{L}_{\text{sim}}(\cdot,\cdot)$ is the *similarity measure* and $\mathcal{L}_{\text{reg}}(\cdot)$ is a *regularizer* with $\lambda \geq 0$.

Diffeomorphic transformations. In medical image registration it is frequently desirable to obtain diffeomorphic transformation maps, which are desirable because they allow mapping smoothly between image spaces (e.g., between an image and an atlas space) without worrying about transformation folds. As discussed above, the standard approach to obtain diffeomorphic transformations by construction is by parameterizing them by velocity [54] or momentum fields [5]. Instead, we will use a regularizer (see Sec. 4.1) that directly encourages invertibility and empirically results in approximately diffeomorphic transformations.

4 Methods

This section describes our approximately diffeomorphic neural deformation fields (NDFs) (Sec. 4.1), how we obtain generalizable neural deformation fields (Sec. 4.2), and lastly our training approach (Sec. 4.3).

4.1 Approximately Diffeomorphic NDFs

We build on the approach of Tian et al. [50] to obtain *approximately* diffeomorphic neural transformation maps. This approach regularizes transformations based on a gradient

[3] A perfect match can in general not be achieved due to image differences or image noise.

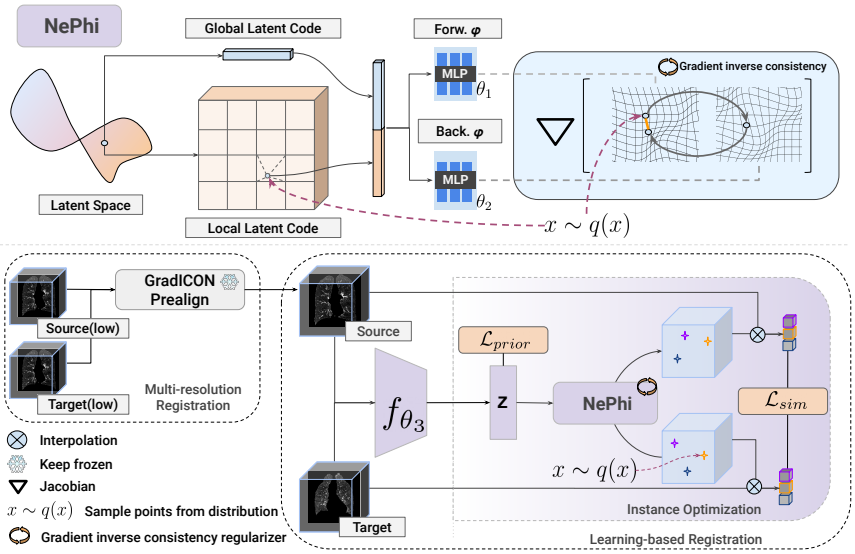

Fig. 2: The framework of `NePhi`. Top: the design of the generalizable neural deformation fields with gradient inverse consistency regularizer (Sec. 4.1 and Sec. 4.2). Bottom: the overall framework of how to train neural networks to predict `NePhi` (Sec. 4.3).

inverse consistency loss to encourage diffeomorphic transforms. Specifically, given a forward transformation map $\varphi^{AB} = \text{Id} + f_\theta(I^A, I^B)$ and a backward map $\varphi^{BA} = \text{Id} + f_\theta(I^B, I^A)$, where Id denotes the identity map, regularity can be encouraged by minimizing the loss

$$\mathcal{L}_{reg} = \left\| \nabla \left[\varphi^{AB} \circ \varphi^{BA} \right] - \mathbf{I} \right\|_F^2 , \tag{2}$$

where $\mathbf{I}$ is the identity matrix and ∇ denotes the Jacobian. To combine this loss with our neural representation of `NePhi`, we use *two* MLPs, denoted φ_{θ_1} and φ_{θ_2}, to represent the continuous forward transformation φ^{AB} and backward transformation φ^{BA}, respectively. To make these MLPs generalizable we condition them on the *same* latent code z, which is possible because of our design choice to use two MLPs. With such a design, we can then express the gradient inverse consistency loss as

$$\mathcal{L}_{reg}(\theta_1, \theta_2, z) = E_{x \sim q(x)} \left\| \nabla \left[\varphi_{\theta_1}(\varphi_{\theta_2}(x; z); z) - x \right] \right\|_F^2 , \tag{3}$$

where x denotes a point coordinate which we uniformly sample from distribution $q(x)$ across the image domain.

4.2 Hybrid Conditioning

To allow `NePhi` to generalize to new transformations (for new image pairs) we use the latent code, z, introduced above to condition φ_{θ_1} and φ_{θ_2}. Specifically, z is composed of a global latent code *vector* $z_g \in \mathbb{R}^{C \times 1}$ and a local latent code *map* $z_l \in \mathbb{R}^{C \times H \times W}$ for a 2D transformation and $z_l \in \mathbb{R}^{C \times D \times H \times W}$ for a 3D transformation. Note that the local

latent code is represented at a lower resolution than the input images. For any point x we then obtain a local latent code via bilinear (trilinear in 3D) interpolation of the local latent codes. We denote the interpolated local latent code at point x by $z_l(x) \in \mathbb{R}^{C\times 1}$. For a given x, the overall latent code, $z \in \mathbb{R}^{2C\times 1}$ is then obtained by concatenation

$$z(x) = [z_g^T, z_l(x)^T]^T \,. \tag{4}$$

We use the following loss from [32] to encourage a compact latent space

$$\mathcal{L}_{prior}(z_g, z_l) = \frac{1}{\sigma^2}||z_g||_2^2 + \frac{1}{|\Omega_{z_l}|}\sum_{x\in\Omega_{z_l}} \frac{1}{\sigma^2}||z_l(x)||_2^2 \tag{5}$$

where Ω_{z_l} represents the grids where the z_l are defined.

4.3 Training

Here, we describe how NePhi can be used for learning-based registration. Fig. 2 (bottom) illustrates the approach. Supp. A provides implementation details.

Learning-based NePhi. We condition NePhi on latent codes because we want NePhi to provide equivalent functionality to the voxel representation used in learning-based registration. In a learning-based registration network, a CNN directly predicts the transformation given a pair of images during inference. In NePhi, we achieve a similar functionality by letting a neural network f_{θ_3} predict the latent codes z_g and z_l based on an input image pair. Given a training dataset $I = \{(I_i^A, I_i^B)\}_{i=1}^N$, we train the neural network f_{θ_3} and NDFs φ_{θ_1}, φ_{θ_2} by minimizing

$$\begin{aligned}\{\theta_1,\theta_2,\theta_3\} = \underset{\theta_1,\theta_2,\theta_3}{\operatorname{argmin}} \frac{1}{N}\sum_{i=1}^{N} \mathcal{L}_{\text{sim}}\left(I_i^A \circ \varphi_{\theta_1}(\cdot;z), I_i^B\right) + \mathcal{L}_{\text{sim}}\left(I_i^A, I_i^B \circ \varphi_{\theta_2}(\cdot;z)\right) \\ + \lambda_1\mathcal{L}_{\text{reg}}(\theta_1,\theta_2,z) + \lambda_2\mathcal{L}_{prior}(z)\,, z = f_{\theta_3}(I_i^A, I_i^B)\,,\end{aligned} \tag{6}$$

where $\mathcal{L}_{sim}$ is the similarity loss: in this work, we use normalized cross correlation (NCC). As φ_{θ_1} and φ_{θ_2} are continuous functions that can represent transformations at any spatial location, we randomly sample two different sets of points within the image space to compute $\mathcal{L}_{\text{sim}}$ and $\mathcal{L}_{\text{reg}}$.

NePhi Instance Optimization. Starting from the predicted latent code from the learning-based NePhi, we optimize over the latent code and the parameters of φ_{θ_1} and φ_{θ_2} by minimizing the following loss at test time when given one pair of images:

$$\begin{aligned}\{z,\theta_1,\theta_2\} = \underset{z,\theta_1,\theta_2}{\operatorname{argmin}} \mathcal{L}_{\text{sim}}\left(I^A \circ \varphi_{\theta_1}(\cdot;z), I^B\right) + \mathcal{L}_{\text{sim}}\left(I^A, I^B \circ \varphi_{\theta_2}(\cdot;z)\right) \\ + \lambda_1\mathcal{L}_{\text{reg}}(\theta_1,\theta_2,z) + \lambda_2\mathcal{L}_{prior}(z)\,.\end{aligned} \tag{7}$$

During training the latent code is predicted by f_{θ_3}. During instance optimization, the latent code is optimized over.

Multi-resolution NePhi To improve inference performance, we propose a multi-resolution approach based on NePhi. We combine a low-resolution voxel-based registration CNN with a high-resolution learning-based **NePhi** as shown in Figure 2. This

approach retains the good memory efficiency of NePhi as the voxel-based registration at low-resolution is also fast and memory efficient. See Fig. 3 for an illustration of NePhi's memory efficiency for high resolution images. Specifically, we train a voxel-based registration neural network on low-resolution images and keep it frozen while training a subsequent high-resolution registration step based on NePhi.

Sampling Strategy As we will observe in Sec. 5.6, regular transforms can be obtained using very few sampled points for the regularizer. Thus, we sample two different sets of points uniformly within the image domain $[-1, 1]$ to compute $\mathcal{L}_{\text{sim}}$ and $\mathcal{L}_{\text{reg}}$, with a much smaller number of points for $\mathcal{L}_{\text{reg}}$ than for $\mathcal{L}_{\text{sim}}$.

5 Experiments

5.1 Datasets

We use the following datasets for our experiments.

Triangles and Circles (T&C) is a synthetic 2D dataset [16] where images are either solid or hollow triangles or circles. We randomly generate a training set of 10000 pairs of images of size 512×512 and a validation set with 1000 pairs of images of the same size for solid and for hollow shapes. We use this dataset for an ablation study.

COPDGene [37]. We use a subset of 999 inspiration/expiration lung computed tomography (CT) image pairs from the **COPDGene** study[4] [37] with provided lung segmentation masks for training. We resample the images to an isotropic spacing of 2 mm, leading to a $175 \times 175 \times 175$ image grid. CT intensities correspond to Hounsfield units. We clamp them within $[-1000, 0]$ and scale them linearly to $[0, 1]$. We then apply the lung segmentation mask to the images to extract the lung region of interest (ROI). We use 899 pairs for training and 100 pairs for validation.

DirLab [6]. This dataset contains 10 pairs of inspiration/expiration lung CT images (from COPDGene, different from the 999 pairs above) with 300 anatomical landmarks per pair, manually identified by an expert in thoracic imaging. We process these images in the same way as the other **COPDGene** images. This dataset is used for testing.

HCP [52]. We use a subset of the young adult T1-weighted brain images of the Human Connectome Project (HCP) for training. We use 28 manually segmented subcortical regions for evaluation [38, 39]. Following [50] we use 1076 images for training and 44 for testing.

5.2 Baselines

We compare NePhi to the following optimization-based method and two learning-based methods.

PTVReg [55] is a state-of-the-art optimization-based approach on the DirLab dataset. We report the result on DirLab based on the official implementation, where hyper-parameters were obtained by grid-search on the DirLab dataset itself. Results are therefore likely overly optimistic.

[4] https://www.ncbi.nlm.nih.gov/gap

LapIRN [30] is the winner of the Learn2Reg registration challenge. To be consistent with the transformation model of NePhi, we use LapIRN with displacement vector fields (DVFs) from the official implementation and train it with the COPDGene dataset.

GradICON [50] is currently the best learning-based registration method on the DirLab dataset. It is a multi-resolution registration neural network composed of three sub-networks. Here a subnetwork receives as input the warped images from the previous network. We compare to GradICON in two ways: 1) based on the original settings, and 2) by replacing the multi-resolution network architecture with a single-resolution architecture. All other settings are kept the same as described in [50]. The goal of the single-resolution comparision is to provide a fair comparison to a single-resolution NePhi model. For GradICON(IO), we use 50 iterations as in the original paper.

Table 1: Registration performance on **DirLab**. Registration methods are categorized as optimization-based (O), multi-resolution learning-based (M), and single-resolution learning-based (S). mTRE is the landmark error. $\%|J|_{<0}$ is the voxel percentage with negative Jacobian determinant for the transformation map, which measures the regularity of the transformation. Rep. shows the number of parameters of a representation and Enc. the number of parameter of the encoder network. The inference time for IO indicates the time required for the optimization iterations. Our NePhi approach combines good performance with excellent registration regularity while reducing memory consumption significantly in comparison to the fully voxel-based GradICON and LapIRN registration networks.

	Method	IO	mTRE↓ (mm)	$\%\|J\|_{<0}\downarrow$	Rep. (M)	Enc. (M)	Time(Inference) (s)	Peak Mem.(Train)↓ (MB)
	Initial		23.36	—	—	—	—	—
O	PTVReg [55]		0.84	0.6	—	—	442	—
S	GradICON		5.41	1.4e-4	0	17.67	0.14	6394
S	GradICON	✓	2.10	3.5e-4	0	17.67	16.08	6394
S	NePhi		5.44	0.0	0.28	3.07	0.46	**2466**
S	NePhi	✓	1.73	0	0.28	0	13.60	2330
M	LapIRN		4.24	1.1e-2	0	0.92	0.24	10186
M	GradICON		1.93	2.6e-4	0	70.68	0.16	13482
M	GradICON	✓	1.31	2.6e-4	0	70.68	47.63	13482
M	NePhi		2.78	2.4e-4	0.28	3.07	0.52	**2852**
M	NePhi	✓	1.53	1.1e-4	0.28	0	56.42	2330

5.3 Registration Performance on DIRLab and HCP

Table 1 compares NePhi to baselines on the DirLab dataset. Note that the iterations of NePhi(IO) are adjusted to match the runtime of GradICON(IO) for a fair comparison.

For single-resolution registration, our method achieves comparable registration accuracy (5.44 mm) to GradICON (5.41 mm), the analogous baseline for a voxel-represented deformation, while using only 41.4% of the memory required for both training and inference. With instance optimization, NePhi with 200 iterations reaches better perfor-

mance than GradICON with 50 iterations (1.73 mm versus 2.10 mm) at a comparable run-time, but again requires less memory. Moreover, the deformation fields predicted by NePhi contain *no* folds, indicating better regularity than GradICON and much better regularity than the optimization-based SOTA method PTVReg.

For multi-resolution registration, NePhi achieves better registration accuracy and regularity than LapIRN. Though the inference result of NePhi is slightly worse than for multi-resolution GradICON, memory consumption is reduced by $\sim 79\%$ compared to GradICON. This memory efficiency carries over to instance optimization, where the performance gap between NePhi(IO) and GradICON(IO) is further reduced for the same runtime. See Fig. 6 in the supplementary material for an example registration result.

Table 2: Evaluating the multi-resolution NePhi on the HCP brain dataset. As for the DirLab dataset, NePhi provides competitive registration accuracy with excellent spatial regularity and significantly reduced memory requirements compared to the voxel-based GradICON registration network during training.

	DICE↑	$\%\lvert J\rvert_{<0}\downarrow$	PeakMem.(Train)↓
Initial	53.4	—	—
GradICON	78.7	1.2e-3	6716
GradICON(IO)	80.5	4e-4	6716
NePhi	76.2	0	**1760**
NePhi(IO)	78.9	1.4e-3	**1426**

We also test NePhi on the HCP brain MRI dataset following the same evaluation procedure as in [50]. Tab. 2 indicates that NePhi shows competitive performance with significantly reduced memory requirements.

5.4 Memory consumption across image resolutions

In the previous experiment, we demonstrated that NePhi exhibits superior memory efficiency compared to state-of-the-art (SOTA) learning-based registration methods for images of size $175 \times 175 \times 175$. Here, we extend our analysis to show that the memory efficiency of NePhi becomes more pronounced as the input image resolution increases.

Figure 3 compares peak memory use between GradICON and NePhi during training and instance optimization (IO) in the single-resolution setting. The results show that peak memory use escalates rapidly for voxel-represented transformations as the size of the input images increases. We see that going from 160^3 to 320^3, memory use for a voxel–based approach jumps from ~ 4.8GB to ~ 37.7GB with only one pair of images and hits the GPU memory limit for a 48 GB GPU around 360^3. In contrast, NePhi shows a significantly slower increase in peak memory use opening the possibility for 3D registrations for significantly larger image volumes. The peak memory use for NePhi is further reduced because optimization is carried out *only* over the parameters of the neural deformation fields, eliminating the need to optimize over the feature encoder. In

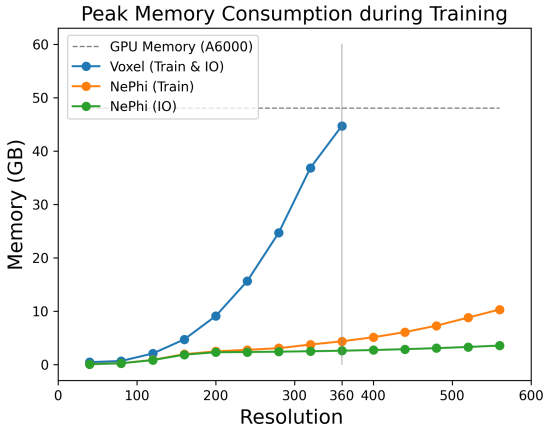

Fig. 3: Comparison of peak memory consumption between voxel-representated deformations and NePhi during training and instance optimization (test-time optimization) for 3D images. Details regarding the experiment setting can be found in Sec. 5.4. We observe that memory consumption for voxel-based approaches scales poorly whereas NePhi requires much less memory during training and testing thereby opening opportunities for high-resolution image registration. Such high-resolution capabilities are expected to become increasingly important for example for the registration of very large volumetric miscroscopy images.

contrast, optimizing over voxel-represented transformation fields with gradient inverse consistency as a regularizer (see Supp. B.2), is not feasible. Peak memory consumption during IO remains high for voxel-represented transformation, equaling that of the training phase.

Table 3: Performance of NePhi on **DirLab** when trained on different resolutions of the COPDGene dataset. We observe that regularity is maintained across resolutions while registration performance improves for higher-resolution images.

Shape	mTRE↓	$\%\lvert J\rvert_{<0}\downarrow$	mTRE(IO)↓	$\%\lvert J\rvert_{<0}\downarrow$
21	16.85	0	11.49	0
43	12.77	0	8.20	0
87	9.31	0	3.48	0
175	5.44	0	1.57	3e-5
256	5.30	0	1.46	3e-4

Table 4: mTRE (mm) and run-time with varying number of IO iterations for DirLab. NePhi(M) outperforms NePhi(S) for real-time inference. Both methods reach satisfying results with reasonable run-time when using IO. NePhi(M) shows better performance for fewer iterations.

IO Iterations	Initial	100	200	800	1600
Time(s)		6.76	13.60	56.42	112.68
NePhi(S)	5.44	1.94	1.73	1.57	1.51
NePhi(M)	2.78	1.78	1.67	1.53	1.49

Beyond memory considerations, Table 3 shows the accuracy and regularity properties based on the predicted deformation field for NePhi trained on various image resolutions for the DirLab dataset. We observe that regularity is maintained across all resolutions, and landmark errors decrease as expected with increasing resolution. This analysis underscores the positive correlation between high-resolution input images and improved registration performance, emphasizing the necessity of exploring high-resolution registration.

5.5 Trade-off between time and accuracy

Table 4 shows registration accuracies for instance optimization (IO) using single-resolution `NePhi` and multi-resolution `NePhi` for varying numbers of IO iterations on the DirLab dataset. We observe that using more iterations results in better registration performance. Multi-resolution registration provides slightly worse than SOTA results while providing fast memory-efficient inference (0.46s for single-resolution `NePhi` and 0.52s for multi-resolution `NePhi`). To pursue the best accuracy, either single-resolution or multi-resolution `NePhi`(IO) can be used with reasonable run-time ($\sim 56s$). These results demonstrate that `NePhi` can effectively balance accuracy and run-time based on registration requirements.

5.6 Ablation study

This section investigates the impact of i) the number of points used for $\mathcal{L}_{\text{reg}}$, and ii) the different choices of $\mathcal{L}_{\text{reg}}$ when combined with NePhi NDF.

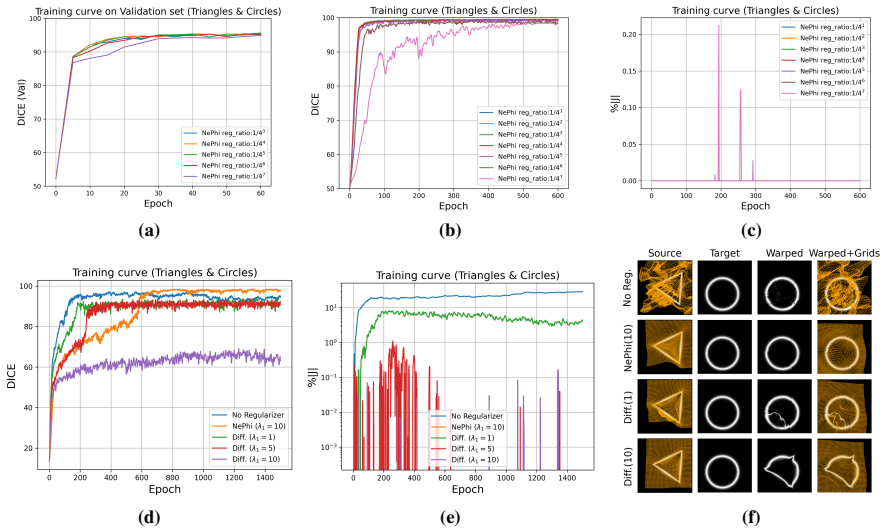

Fig. 4: Ablation study. (a) Performance of `NePhi` trained with various regularizer ratios, indicating that `NePhi`'s performance is insensitive to the regularizer's sampling ratio. (b) and (c) show the registration accuracy and regularity of the transformation field when using `NePhi` in the optimization-based setting (as in IO, but without predicting the latent code). The registration accuracy and regularity of the transformation field only deteriorates for regularizer ratios below $\sim 4^{-6}$. (d), (e) and (f) show comparisons between using the diffusion regularizer and gradient inverse consistency (`NePhi`) on NDF (Neural Deformation Field) transformations in the optimization setting.

The number of points used for $\mathcal{L}_{\text{reg}}$. The underlying principle of NePhi that facilitates memory efficiency lies in its capability to preserve the regularity of deformation fields through the computation of regularity losses over a small number of sparse points, rather than densely across the entire image volume. In this section, we test this hypothesis and investigate the minimal set of points required to preserve regularity. We define the **regularizer ratio** $\%\mathcal{L}_{\text{reg}}$ as the ratio between the number of sample points to evaluate the regularizer loss over the number of image voxels. To eliminate the influence of the similarity loss, we keep the similarity ratio fixed to 0.02 (seeSupp. A for details) for all models in this section.

Varying $\%\mathcal{L}_{\text{reg}}$ in learning-based NePhi. We train learning-based NePhi on the **T&C** solid dataset. Fig. 4a shows the Dice overlap scores between the fixed and the registered moving image. The number of folds ($\%|J|_{<0}$; not shown for this experiment) remained consistently close to zero for all models when varying $\%\mathcal{L}_{\text{reg}}$. Fig. 4a reveals that, at convergence of model training, both registration accuracy and the regularity of predicted deformation fields are minimally affected by variations in the regularizer ratio. Consequently, memory consumption for the evaluation of the regularizer loss can be dramatically reduced without affecting overall registration accuracy and regularity.

Varying $\%\mathcal{L}_{\text{reg}}$ with NePhi(IO). We also tested NePhi(IO) when varying $\%\mathcal{L}_{\text{reg}}$. Learning-based NePhi solves the T&C registration problem well, so performing instance optimization starting from a trained learning-based NePhi results in flat training curves, starting and staying at the optimum. To emphasize convergence differences, we instead initialize the weights of the MLP and the latent code randomly, and then perform instance optimization. We randomly pick one image pair from the **T&C** as an example. Fig. 4b and Fig. 4c show the Dice overlap and transformation regularity respectively. We observe that reducing $\%\mathcal{L}_{\text{reg}}$ does not significantly impact the regularity of the obtained transformations until $\%\mathcal{L}_{\text{reg}} = \frac{1}{4^6} \sim 0.0244\%$ is reached, corresponding to approximately 64 points only ($\%\mathcal{L}_{\text{reg}} \times 512 \times 512$). Importantly, since lowering the regularizer ratio does not affect regularity over a wide range of regularizer ratios, registration accuracy remains unaffected for these settings as well. This experiment underscores that, similar to learning-based NePhi, in this IO setting, satisfactory accuracy can be achieved by evaluating regularity using very few sample points for the regularizer loss, allowing for dramatic reductions in memory use for the evaluation of the regularizer loss.

The different choices of $\mathcal{L}_{\text{reg}}$. We opt for gradient inverse consistency (GradICON) regularization instead of the more commonly used diffusion regularizer [4] for deformation vector fields (DVF). This choice is motivated by the excellent registration performance obtained using the GradICON regularizer for voxel-based learning-based registration and its ability to provide approximately diffeomoprhic transformations when combined with a deep registration network [50]. However, gradient inverse consistency does not work in the optimization-based setting for a voxel-represented transformation (shown in Supp B.2). However, it does work for NePhi, enabling optimization over the transformation field itself instead of the entire network in the IO phase. This immediately also raises the question if gradient inverse consistency results in better solutions comparing to diffusion regularization for NDFs for instance optimization. To answer

this question, we compare NePhi in the IO setting: i) NePhi without using a regularizer, and ii) NePhi replacing gradient inverse consistency with diffusion regularization. As discussed in Sec. 5.6, to emphasize the convergence differences, we initialize NePhi (including latent codes and MLPs) from a random initialization. We present the result on one randomly generated pair of images from the **T&C** hollow dataset (more examples in Supp. B.3). Experimental results are shown in Fig. 4d, Fig. 4e and Fig. 4f. Fig. 4e reveals that NePhi is the only approach that results in zero foldings during optimization. Although the number of folds ($\%|J|_{<0}$) also goes to zero for the models using the diffusion regularizer, the Dice results at convergence are lower compared to NePhi in Fig. 4d. Fig. 4d shows visualizations of the final registration results illustrating the good regularity properties of NePhi based on gradient inverse consistency regularization.

6 Limitations

Our method provides memory benefits during training and testing and allows graceful tradeoff between accuracy and runtime. However, it has the following limitations. First, the convergence is relatively slow. NePhi requires more training epochs (Supp. A) and more iterations (800) during instance optimization to converge to a similar performance level as voxel-based registration neural networks (50 IO iterations are sufficient). This might be addressable by sampling more points to evaluate the losses which would increase memory requirements. Second, the current multi-resolution NePhi is still based on an initial voxel-based CNN registration model. For very high-resolution registration, such low resolution voxel-based networks may themselve consume a significant memory.

7 Conclusion

In this work, we introduced NePhi, which provides comparable registration accuracy, better deformation regularity, flexible settings to balance between test runtime and registration accuracy, and, importantly requires significantly less memory during training and instance optimization, compared to the current SOTA learning-based registration methods. NePhi opens up a large design space for learning-based registration (see Fig. 1) allowing a user to pick the best suited NePhi model for a particular registration task. In this work we demonstrated the good performance of NePhi within this design space on two challenging 3D image registration tasks: lung registration based on the compute tomography images of the COPDGene dataset and brain registration based on the magnetic resonance images of the HCP dataset. We hope this work will inspire further explorations of generalizable neural deformation fields in learning-based registration methods and that it will open avenues for further research into high-resolution registration.

Acknowledgements

The research reported in this publication was supported by the National Institutes of Health (NIH) under award numbers NIH 1R01HL149877, 1R01EB028283, 1R01AR082684, 1R21MH132982, RF1MH126732, and 1R21EB035832. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The brain imaging data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The lung imaging data was provided by the COPDGene study.

References

1. Ashburner, J.: A fast diffeomorphic image registration algorithm. NeuroImage **38**(1), 95–113 (2007)
2. Avants, B., Tustison, N., Song, G.: Advanced normalization tools (ANTS). Insight J **1–35** (11 2008). https://doi.org/10.54294/uvnhin
3. Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J., Dalca, A.V.: An unsupervised learning model for deformable medical image registration. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 9252–9260 (2018)
4. Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J.V., Dalca, A.V.: Voxelmorph: A learning framework for deformable medical image registration. IEEE Transactions on Medical Imaging **38**(8), 1788–1800 (2019)
5. Beg, M.F., Miller, M.I., Trouvé, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. IJCV **61**(2), 139–157 (2005)
6. 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. Phys. Med. Biol. **58**(9), 2861 (2013)
7. Chen, J., Frey, E.C., He, Y., Segars, W.P., Li, Y., Du, Y.: Transmorph: Transformer for unsupervised medical image registration. Medical image analysis **82**, 102615 (2022)
8. Chen, X., Zheng, Y., Black, M.J., Hilliges, O., Geiger, A.: Snarf: Differentiable forward skinning for animating non-rigid neural implicit shapes. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 11594–11604 (2021)
9. Christensen, G.E., Rabbitt, R.D., Miller, M.I.: 3D brain mapping using a deformable neuroanatomy. Phys. Med. Biol. **39**(3), 609 (1994)
10. Christensen, G.E., Rabbitt, R.D., Miller, M.I.: Deformable templates using large deformation kinematics. TMI **5**(10), 1435–1447 (1996)
11. Crum, W.R., Hartkens, T., Hill, D.: Non-rigid image registration: theory and practice. Brit. J. Radiol. **77**(suppl_2), S140–S153 (2004)
12. Dalca, A., Balakrishnan, G., Guttag, J., Sabuncu, M.: Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces. Medical Image Analysis **57**, 226–236 (2019)
13. Duggal, S., Pathak, D.: Topologically-aware deformation fields for single-view 3D reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1536–1546 (2022)
14. Gao, C., Saraf, A., Kopf, J., Huang, J.B.: Dynamic view synthesis from dynamic monocular video. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 5712–5721 (2021)
15. Grassal, P.W., Prinzler, M., Leistner, T., Rother, C., Nießner, M., Thies, J.: Neural head avatars from monocular RGB videos. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 18653–18664 (2022)
16. Greer, H., Kwitt, R., Vialard, F.X., Niethammer, M.: ICON: Learning regular maps through inverse consistency. In: ICCV (2021)
17. 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. pp. 1869–1879 (2023)
18. van Harten, L., Van Herten, R.L.M., Stoker, J., Isgum, I.: Deformable image registration with geometry-informed implicit neural representations. In: Medical Imaging with Deep Learning (2023)
19. Heinrich, M.P., Jenkinson, M., Brady, M., Schnabel, J.A.: Globally optimal deformable registration on a minimum spanning tree using dense displacement sampling. In: MICCAI. vol. 7512, pp. 115–122 (2012)

20. 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)
21. Heinrich, M.P., Papiez, B.W., Schnabel, J.A., Handels, H.: Non-parametric discrete registration with convex optimisation. In: Biomedical Image Registration - 6th International Workshop, WBIR. vol. 8545, pp. 51–61 (2014)
22. Holden, M.: A review of geometric transformations for nonrigid body registration. TMI **27**(1), 111–128 (2007)
23. Lei, J., Daniilidis, K.: Cadex: Learning canonical deformation coordinate space for dynamic surface representation via neural homeomorphism. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 6624–6634 (2022)
24. Li, Z., Niklaus, S., Snavely, N., Wang, O.: Neural scene flow fields for space-time view synthesis of dynamic scenes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 6498–6508 (2021)
25. Liu, J.W., Cao, Y.P., Mao, W., Zhang, W., Zhang, D.J., Keppo, J., Shan, Y., Qie, X., Shou, M.Z.: Devrf: Fast deformable voxel radiance fields for dynamic scenes. arXiv preprint arXiv:2205.15723 (2022)
26. Liu, L., Habermann, M., Rudnev, V., Sarkar, K., Gu, J., Theobalt, C.: Neural actor: Neural free-view synthesis of human actors with pose control. ACM Transactions on Graphics (TOG) **40**(6), 1–16 (2021)
27. Maes, F., Vandermeulen, D., Suetens, P.: Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. Medical Image Analysis **3**(4), 373–386 (1999). https://doi.org/https://doi.org/10.1016/S1361-8415(99)80030-9, https://www.sciencedirect.com/science/article/pii/S1361841599800309
28. Mehta, I., Gharbi, M., Barnes, C., Shechtman, E., Ramamoorthi, R., Chandraker, M.: Modulated periodic activations for generalizable local functional representations. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 14214–14223 (2021)
29. Mok, T.C.W., Chung, A.C.S.: Fast symmetric diffeomorphic image registration with convolutional neural networks. In: IEEE CVPR. pp. 4643–4652 (2020)
30. Mok, T.C.W., Chung, A.C.S.: Large deformation diffeomorphic image registration with Laplacian pyramid networks. In: Medical Image Computing and Computer Assisted Intervention. vol. 12263, pp. 211–221 (2020)
31. Niemeyer, M., Mescheder, L., Oechsle, M., Geiger, A.: Occupancy flow: 4d reconstruction by learning particle dynamics. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 5379–5389 (2019)
32. Park, J.J., Florence, P., Straub, J., Newcombe, R., Lovegrove, S.: Deepsdf: Learning continuous signed distance functions for shape representation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 165–174 (2019)
33. Park, K., Sinha, U., Barron, J.T., Bouaziz, S., Goldman, D.B., Seitz, S.M., Martin-Brualla, R.: Nerfies: Deformable neural radiance fields. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 5865–5874 (2021)
34. Park, K., Sinha, U., Hedman, P., Barron, J.T., Bouaziz, S., Goldman, D.B., Martin-Brualla, R., Seitz, S.M.: Hypernerf: A higher-dimensional representation for topologically varying neural radiance fields. arXiv preprint arXiv:2106.13228 (2021)
35. Peng, S., Dong, J., Wang, Q., Zhang, S., Shuai, Q., Zhou, X., Bao, H.: Animatable neural radiance fields for modeling dynamic human bodies. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 14314–14323 (2021)
36. Pumarola, A., Corona, E., Pons-Moll, G., Moreno-Noguer, F.: D-nerf: Neural radiance fields for dynamic scenes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 10318–10327 (2021)

37. Regan, E.A., Hokanson, J.E., Murphy, J.R., Make, B., Lynch, D.A., Beaty, T.H., Curran-Everett, D., Silverman, E.K., Crapo, J.D.: Genetic epidemiology of COPD (COPDGene) study design. COPD: J. Chronic Obstr. Pulm. Dis. **7**(1), 32–43 (2011)
38. Rushmore, R.J., Sutherland, K., Carrington, H., Chen, J., Halle, M., Lasso, A., Papadimitriou, G., Prunier, N., Rizzoni, E., Vessey, B., Wilson-Braun, P., Rathi, Y., Kubicki, M., Bouix, S., Yeterian, E., Makris, N.: Anatomically curated segmentation of human subcortical structures in high resolution magnetic resonance imaging: An open science approach. Front. Neuroanat. **16** (2022)
39. Rushmore, R.J., Sutherland, K., Carrington, H., Chen, J., Halle, M., Lasso, A., Papadimitriou, G., Prunier, N., Rizzoni, E., Vessey, B., Wilson-Braun, P., Rathi, Y., Kubicki, M., Bouix, S., Yeterian, E., Makris, N.: HOA-2/SubcorticalParcellations: release-50-subjects-1.1.0 (Sep 2022). https://doi.org/10.5281/zenodo.7080547, https://doi.org/10.5281/zenodo.7080547
40. Shao, R., Zhang, H., Zhang, H., Chen, M., Cao, Y.P., Yu, T., Liu, Y.: Doublefield: Bridging the neural surface and radiance fields for high-fidelity human reconstruction and rendering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 15872–15882 (2022)
41. Shen, Z., Han, X., Xu, Z., Niethammer, M.: Networks for joint affine and non-parametric image registration. In: IEEE CVPR. pp. 4224–4233 (2019)
42. Siebert, H., Hansen, L., Heinrich, M.P.: Fast 3D registration with accurate optimisation and little learning for Learn2Reg 2021. In: Biomedical Image Registration, Domain Generalisation and Out-of-Distribution Analysis. vol. 13166, pp. 174–179 (2021)
43. Sitzmann, V., Martel, J., Bergman, A., Lindell, D., Wetzstein, G.: Implicit neural representations with periodic activation functions. Advances in Neural Information Processing Systems **33** (2020)
44. Sotiras, A., Davatzikos, C., Paragios, N.: Deformable medical image registration: A survey. TMI **32**(7), 1153–1190 (2013)
45. Sun, S., Han, K., Kong, D., You, C., Xie, X.: Mirnf: medical image registration via neural fields. arXiv preprint arXiv:2206.03111 (2022)
46. Sun, W., Niessen, W.J., Klein, S.: Free-form deformation using lower-order B-spline for nonrigid image registration. In: MICCAI. pp. 194–201 (2014)
47. Sundararaman, R.S., Marin, R., Rodola, E., Ovsjanikov, M.: Reduced representation of deformation fields for effective non-rigid shape matching. Advances in Neural Information Processing Systems **35**, 10405–10420 (2022)
48. Tancik, M., Srinivasan, P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J., Ng, R.: Fourier features let networks learn high frequency functions in low dimensional domains. Advances in Neural Information Processing Systems **33**, 7537–7547 (2020)
49. Thirion, J.P.: Image matching as a diffusion process: an analogy with Maxwell's demons. MedIA **2**(3), 243–260 (1998)
50. Tian, L., Greer, H., Vialard, F.X., Kwitt, R., Estépar, R.S.J., Niethammer, M.: GradICON: Approximate diffeomorphisms via gradient inverse consistency. arXiv preprint arXiv:2206.05897 (2022)
51. Tretschk, E., Tewari, A., Golyanik, V., Zollhöfer, M., Lassner, C., Theobalt, C.: Non-rigid neural radiance fields: Reconstruction and novel view synthesis of a dynamic scene from monocular video. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 12959–12970 (2021)
52. Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T.E., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S.W., Della Penna, S., Feinberg, D., Glasser, M.F., Harel, N., Heath, A.C., Larson-Prior, L., Marcus, D., Michalareas, G., Moeller, S., Oostenfeld, R.,

Peterson, S.E., Prior, F., Schlaggar, B.L., Smith, S.M., Snyder, A.Z., Xu, J., Yacoub, E., Consortium, W.M.H.: The Human Connectome Project: a data acquisition perspective. Neuroimage **62**(4), 2222–2231 (2012)
53. Van Harten, L.D., Stoker, J., Išgum, I.: Robust deformable image registration using cycle-consistent implicit representations. IEEE Transactions on Medical Imaging (2023)
54. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Symmetric log-domain diffeomorphic registration: A Demons-based approach. In: MICCAI (2008)
55. 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)
56. Wang, Z., Deng, Y., Yang, J., Yu, J., Tong, X.: Generative deformable radiance fields for disentangled image synthesis of topology-varying objects. arXiv preprint arXiv:2209.04183 (2022)
57. 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)
58. Wu, N., Zhang, M.: NeurEPDiff: Neural operators to predict geodesics in deformation spaces. arXiv preprint arXiv:2303.07115 (2023)
59. Xu, H., Alldieck, T., Sminchisescu, C.: H-nerf: Neural radiance fields for rendering and temporal reconstruction of humans in motion. Advances in Neural Information Processing Systems **34**, 14955–14966 (2021)
60. Yang, X., Kwitt, R., Styner, M., Niethammer, M.: Quicksilver: Fast predictive image registration–a deep learning approach. NeuroImage **158**, 378–396 (2017)
61. Zhang, R., Chen, J.: NDF: Neural deformable fields for dynamic human modelling. In: Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXXII. pp. 37–52. Springer (2022)
62. Zheng, Y., Abrevaya, V.F., Bühler, M.C., Chen, X., Black, M.J., Hilliges, O.: Im avatar: Implicit morphable head avatars from videos. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 13545–13555 (2022)
63. Zou, J., Debroux, N., Liu, L., Qin, J., Schönlieb, C.B., Aviles-Rivero, A.I.: Homeomorphic image registration via conformal-invariant hyperelastic regularisation. arXiv preprint arXiv:2303.08113 (2023)