

# Learning Geodesics of Geometric Shape Deformations From Images

Nian Wu <sup>1</sup> , Miaomiao Zhang <sup>1,2</sup> 

<sup>1</sup> University of Virginia, Department of Electrical & Computer Engineering, Charlottesville, VA, USA

<sup>2</sup> University of Virginia, Department of Computer Science, Charlottesville, VA, USA

## Abstract

This paper presents a novel method, named geodesic deformable networks (GDN), that for the first time enables the learning of geodesic flows of deformation fields derived from images. In particular, the capability of our proposed GDN being able to predict geodesics is important for quantifying and comparing deformable shape presented in images. The geodesic deformations, also known as optimal transformations that align pairwise images, are often parameterized by a time sequence of smooth vector fields governed by nonlinear differential equations. A bountiful literature has been focusing on learning the initial conditions (e.g., initial velocity fields) based on registration networks. However, the definition of geodesics central to deformation-based shape analysis is blind to the networks. To address this problem, we carefully develop an efficient neural operator to treat the geodesics as unknown mapping functions learned from the latent deformation spaces. A composition of integral operators and smooth activation functions is then formulated to effectively approximate such mappings. In contrast to previous works, our GDN jointly optimizes a newly defined geodesic loss, which adds additional benefits to promote the network regularizability and generalizability. We demonstrate the effectiveness of GDN on both 2D synthetic data and 3D real brain magnetic resonance imaging (MRI). Our code is publicly available at <https://github.com/nellie689/GDN>.

## Keywords

Geodesic deformations, Diffeomorphisms, Neural Operator

## Article informations

<https://doi.org/10.59275/j.melba.2025-f9f4>

Volume 3, Received: 2024-10, Published 2025-12

Corresponding author: [bsw3ac@virginia.edu](mailto:bsw3ac@virginia.edu)

©2025 Nian Wu and Miaomiao Zhang. License: CC-BY 4.0



## 1. Introduction

Deformable shape provides important features of objects presented in images, for example, abnormal shape changes of anatomical brain structures are shown to be predictors of neurodegenerative disorders (Qiu et al., 2008; Lorenzi et al., 2010; Wang et al., 2007; Wang and Zhang, 2022). Existing methods have studied a variety of shape representations, including deformation-based descriptors that focus on highly detailed geometric information on dense images (Christensen et al., 1993; Rueckert et al., 2003). With the underlying assumption that objects in many generic classes can be described as deformed versions of an ideal template, descriptors in this class arise naturally by aligning the template to an input image. The resulting transformation between images is then considered as a shape descriptor that reflects local geometric changes. A rich body of literature has been dedicated to harnessing such shape

information for enhanced performance in image analysis, particularly in specialized domains of medical image analysis (Bao et al., 2019; Beekman et al., 2022; Sun et al., 2022) and computational anatomy (Wang et al., 2007; Qiu et al., 2008; Grenander and Miller, 1998). The realm of related research involves a suite of tools designed for comparing, matching, and modeling shapes based on distance metrics in deformation spaces (Qiu et al., 2008; Hong et al., 2017; Niethammer et al., 2019; Qiu et al., 2012). A *geodesic* (a.k.a. a curve whose tangent vector is parallel transported along itself) locally minimizes the Riemannian distance between points on a manifold. Such geodesics can quantify the geometric similarity between objects and serve as a solid mathematical foundation for deformable shape analysis, including regression (Hong et al., 2017; Niethammer et al., 2011), longitudinal analysis (Singh et al., 2013; Hong et al., 2019), and group comparisons (Miller, 2004; Qiu

et al., 2008). Such a metric was first developed in the large deformation diffeomorphic metric mapping (LDDMM) algorithm (Beg et al., 2005; Vialard et al., 2012a), giving rise to a variational principle that expresses the optimal deformation as a geodesic flow of diffeomorphisms (i.e., bijective, smooth, and invertible smooth mappings). The geodesic path of deforming one image to another also serves to regularize the smoothness of the transformations grids with an intact topology of objects.

Recent advances in deep registration networks have achieved remarkable success in learning diffeomorphic transformations between images (Yang et al., 2017; Balakrishnan et al., 2019; Wang and Zhang, 2020; Chen et al., 2022; Kim et al., 2022). These learning-based approaches characterize deformation fields using a time sequence of smooth vector fields governed by nonlinear differential equations. In contrast to traditional registration methods such as LDDMM that rely on iterative optimization schemes (Beg et al., 2005; Vialard et al., 2012a), learning-based methods offer rapid predictions without the need for adhoc manual parameter tuning during the testing phase. Despite the aforementioned advantages, current registration networks primarily focus on learning the initial conditions of deformation models, i.e., initial velocity fields (Yang et al., 2017; Wang and Zhang, 2020; Hinkle et al., 2018). The final deformation field is then obtained by solving a set of differential equations, which can be computationally expensive in high-dimensional image space during testing. More importantly, the definition of geodesic metrics in the deformation spaces, which is critical for quantifying and analyzing shape changes, is not explicitly considered or overlooked in the network learning process. This limits the interpretability and regularizability of the transformation process. Additionally, it may negatively impact the model generalizability by elevating the risk of overfitting the registration loss during training phase.

Recent research on neural operators have been explored to learn the solution of ordinary /partial differential equations (O/PDEs) (Lu et al., 2019; Bhattacharya et al., 2021; Li et al., 2020b,a). Among these, a notable work is the Fourier neural operator (FNO) (Li et al., 2020a), which achieves an inference time orders of magnitude faster than traditional numerical solutions to the fluid flow Navier-Stokes equations (Constantin, 1988), without sacrificing accuracy. Inspired by the work of (Li et al., 2020a), a research group had an initial attempt to utilize Fourier neural operators as a surrogate to learn the solution to geodesic flows governed by the well-known Euler-Poincaré differential equation (EPDiff) (Arnold, 1966; Miller et al., 2006) in the context of optimization-based image registration (Wu and Zhang, 2023). Similarly, a group of research work formulate image registration as learning a transformation flow between images, where the deformation trajectories are obtained by

integrating time-dependent velocity fields approximated by neural networks within a neural ODE framework (Sun et al., 2024; Wu et al., 2022). However, they still need a conventional iterative optimization-based scheme for estimating the initial velocity fields during the registration inference phase.

In this paper, we introduce a novel geodesic deformable network (GDN) that *for the first time learns the geodesic metrics of latent deformation models directly from training images* in the context of registration networks. Our method builds upon FNO Li et al. (2023) to model geodesic learning in a latent space, enabling efficient approximation of complex deformation dynamics. Unlike existing approaches (Sun et al., 2024; Wu et al., 2022), our proposed method takes a fundamentally different approach by leveraging the EPDiff geodesic shooting equation to guide the learning of time-dependent velocity fields constrained along geodesic paths. In comparison to the recent NeurEPDiff method (Wu and Zhang, 2023), our model differs by treating the geodesics as unknown mapping functions in the latent deformation spaces, which are explicitly learned in the context of deep registration networks, without requiring optimization during the registration inference stage. A composition of integral operators and smooth activation functions is then formulated to effectively approximate such mappings in each hidden layer of the proposed GDNs. Our network will jointly optimize a newly defined geodesic loss, guided by the underlying transformation process solved through numerical integrators (Vialard et al., 2012a; Beg et al., 2005). A major benefit of the learned geodesic mapping function lies in its application as a computationally efficient surrogate model for the original numerical solutions of differential equations required to generate final deformations. In contrast to previous work (Balakrishnan et al., 2019; Wang and Zhang, 2020; Chen et al., 2022; Kim et al., 2022), the advantages of our proposed method, GDN, are threefold:

- Predict a geodesic flow of deformation fields using a surrogate model, which is important for deep nets to quantify and analyze deformation-based shape objects.
- Promote the network regularizability and generalizability via enforced geodesic constraints in deformation spaces.
- Eliminate the need for numerical solutions of differential equations for final deformations in testing inference, a strategy that can be computationally advantageous in high-dimensional image spaces.

Experiments on 2D synthetic data and 3D real brain MRIs show that GDN is able to predict fairly close geodesics compared to conventional numerical solutions. We also demonstrate that GDN achieves superior performance in regularizability and generalizability when compared to the state-of-the-art deep learning-based registration networks in

out-of-distribution (OOD) data (Balakrishnan et al., 2019; Hinkle et al., 2018; Chen et al., 2022; Yang et al., 2017).

## 2. Background: Geodesics In Deformation Spaces

In this section, we first briefly review the basic mathematical concepts of *geodesics* in the space of diffeomorphic transformations (a.k.a. diffeomorphisms). We then show how to derive the geodesic path of transformations from images using the LDDMM algorithm (Beg et al., 2005) with the geodesic shooting equation (Vialard et al., 2012a; Younes et al., 2009).

### 2.1 Geodesics of Diffeomorphisms

Let  $\text{Diff}^\infty(\Omega)$  denote the space of smooth diffeomorphisms on an image domain  $\Omega$ . The tangent space of diffeomorphisms is the space  $V = \mathfrak{X}^\infty(T\Omega)$  of smooth vector fields on  $\Omega$ . Consider a time-varying velocity field,  $\{v_t\} : [0, \tau] \rightarrow V$ , we can then generate diffeomorphisms  $\{\phi_t\}$  as a solution to the equation

$$\frac{d\phi_t}{dt} = v_t(\phi_t), t \in [0, \tau]. \quad (1)$$

The LDDMM (Beg et al., 2005) provides a *distance metric* in the space of diffeomorphisms, which is also used as a regularization for image registration. Such a distance metric is formulated as an integral of the Sobolev norm of the time-dependent velocity field  $v_t$ , i.e.,  $\int_0^\tau (\mathcal{L}v_t, v_t) dt$ , where  $\mathcal{L} : V \rightarrow V^*$  is a symmetric, positive-definite differential operator that maps a tangent vector  $v(t) \in V$  into its dual space as a momentum vector  $m(t) \in V^*$ . We typically write  $m(t) = \mathcal{L}v(t)$ , or  $v(t) = \mathcal{K}m(t)$ , with  $\mathcal{K}$  being an inverse operator of  $\mathcal{L}$ . In this paper, we adopt a commonly used Laplacian operator  $\mathcal{L} = (-\alpha\Delta + \beta \text{Id})^2$ , where  $\alpha$  and  $\beta$  are weighting parameters that control the smoothness of transformation fields and  $\text{Id}$  is an identity matrix. The  $(\cdot, \cdot)$  is a dual pairing, which is similar to an inner product between vectors.

According to a well-known geodesic shooting algorithm (Vialard et al., 2012a), the minimum of the previously defined distance metric above is uniquely determined by solving an Euler-Poincaré differential (EPDiff) equation (Vialard et al., 2012a; Younes et al., 2009) with a given initial condition. That is, for  $\forall v_0 \in V$ , at  $t = 0$ , a geodesic path  $t \mapsto \phi_t \in \text{Diff}^\infty(\Omega)$  in the space of diffeomorphisms can be computed by forward shooting the EPDiff equation formulated as

$$\frac{\partial v_t}{\partial t} = -K \left[ (Dv_t)^T m_t + Dm_t v_t + m_t \text{div} v_t \right], \quad (2)$$

where  $D$  denotes the Jacobian matrix and  $\text{div}$  is a divergence operator.

### 2.2 LDDMM To Derive Geodesics From Images

Assume a source image  $S$  and a target image  $T$  defined on a  $d$ -dimensional torus domain  $\Omega = \mathbb{R}^d / \mathbb{Z}^d$  ( $S(x), T(x) : \Omega \rightarrow \mathbb{R}$ ). The space of diffeomorphisms is denoted by  $\text{Diff}(\Omega)$ . The problem of diffeomorphic image registration is to find the geodesic, to generate time-varying diffeomorphisms  $\{\phi_t\} : t \in [0, \tau]$ , such that a deformed source image by the smooth mapping  $\phi_\tau$ , noted as  $S(\phi_\tau)$ , is similar to  $T$ . By parameterizing the transformation  $\phi_\tau$  with an initial velocity field  $v_0$ , we write the optimization problem of diffeomorphic registration in the setting of LDDMM with geodesic shooting as

$$E(v_0) = (\mathcal{L}v_0, v_0) + \lambda \text{Dist}(S(\phi_\tau), T) \quad \text{s.t. Eq. (1) \& (2)}. \quad (3)$$

Here,  $\text{Dist}(\cdot, \cdot)$  is a distance function that measures the dissimilarity between images and  $\lambda$  is a positive weighting parameter. The commonly used distance functions include the sum-of-squared intensity differences ( $L_2$ -norm) (Beg et al., 2005), normalized cross correlation (NCC) (Avants et al., 2008), and mutual information (MI) (Wells et al., 1996). In this paper, we use the sum-of-squared intensity differences.

## 3. Our Method

In this section, we present a novel method, GDN, that learns geodesics of deformations in the context of deep registration networks. The key component of GDN is a newly designed geodesic neural operator (GNO) that encodes the underlying geodesic mapping functions in the latent feature space. A joint optimization of a typical unsupervised registration loss with a geodesic loss guided by numerical solutions of geodesic shooting equations will be introduced in GDN. An overview of our network GDN is shown in Fig. 1.

### 3.1 Geodesic Neural Operator

Inspired by recent works on neural operators (Kovachki et al., 2021; Li et al., 2020a), we develop GNO, noted by  $\mathcal{G}_{\theta_r}$  with parameters  $\theta_r$ , as a surrogate model to approximate the solution to the geodesic shooting equations (Eq. (2)) in the latent deformation spaces, parameterized by initial velocity fields.

We first employ a widely adopted unsupervised registration architecture featuring a UNet backbone (Balakrishnan et al., 2019; Hinkle et al., 2018; Chen et al., 2022; Kim et al., 2022) to project the input images into a latent deformation space, denoted as  $\mathcal{Z}$ . This space embeds a collection of latent features,  $z_{v_0} \in \mathcal{Z}$ , associated with the initial velocity field.

Our GNO is then formulated as an iterative architecture,  $z_{v_0} \mapsto z_{v_1} \mapsto \dots \mapsto z_{v_\tau}$ , based on recurrent neural networks.

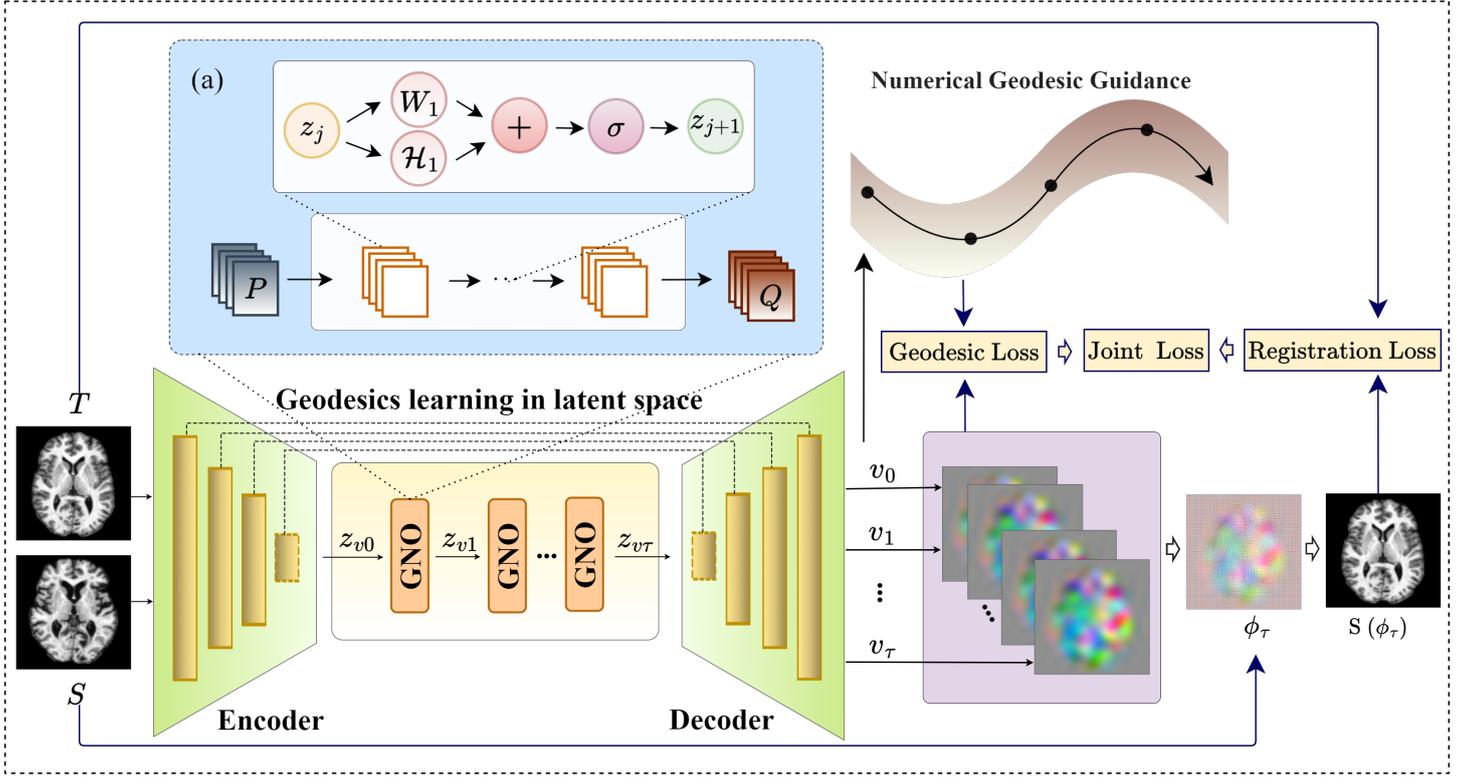


Figure 1: An overview of our proposed network GDN.

As shown in Fig. 1(a), the main architecture of GNO is summarized below:

- An encoder  $P$ : lifts the input latent feature,  $z_{v_0}$ , to a higher dimensional representation with an increased number of channels.
- A *geodesic evolution layer*: learns a nonlinear mapping function of  $z_{v_t} \mapsto z_{v_{t+1}}$  along the geodesic path.
- A decoder  $Q$ : projects the latent mapping function back to the original input dimension aligned with  $z_{v_0}$ .

Analogous to (Li et al., 2020a), the encoder  $P$  and decoder  $Q$  can be local transformations parameterized by a shallow fully connected neural network. Please note that the functionalities of  $P$  and  $Q$  slightly differ from a typical image encoder and decoder, as they expand the latent features,  $z_0$ , into higher-dimensional channels and then project them back to the original channel dimension. This particular design aims to enrich the model capacity and expressivity in the latent space. Next, we will introduce details of the new geodesic evolution layer.

**Geodesic evolution layer.** We utilize a multilayer neural network with  $J$  hidden layers to simulate the geodesic mapping function from  $z_{v_t}$  to  $z_{v_{t+1}}$ . To simplify the math notation, we let  $u^j$  denote the input of  $j$ -th hidden layer and  $u^{j+1}$  denotes the output, where  $j \in \{1, \dots, J\}$ . A

composition of a local linear transformation,  $W^j$ , and a global convolutional kernel,  $\mathcal{H}^j$ , is used to extract both global and local representations at each hidden layer. Additionally, a nonlinear activation function,  $\sigma(\cdot)$ , is carefully designed to encourage a smooth evolution of the geodesic path. We update  $u^j \mapsto u^{j+1}$  by

$$u^{j+1} := \sigma(W^j u^j + \mathcal{H}^j * u^j), \quad (4)$$

where  $*$  represents a convolution operator. In order to compute the global convolution in an efficient way, we apply matrix-vector multiplication for two Fourier signals in the frequency domain, and then project it back to the original space. We employ a Gaussian Error Linear Unit (GELU) (Hendrycks and Gimpel, 2016) with the smoothing operator  $K$  to ensure the smoothness of the output signal, we have  $\sigma := K(\text{GELU}(\cdot))$ .

Finally, we update  $z_{v_{t-1}}$  to  $z_{v_t}$  by

$$z_{v_{t+1}} := \sigma(W^J, \mathcal{H}^J) \circ \dots \circ \sigma(W^2, \mathcal{H}^2) \circ \sigma(W^1, \mathcal{H}^1, z_{v_t}),$$

where  $\circ$  denotes a function composition.

**Geodesic loss.** We formulate the geodesic loss function of GNO as an empirical data loss term guided by numerical solutions (e.g., using an Euler integrator) to the geodesic shooting equations. Since our model GNO is now learning a sequence of latent velocity fields  $\{z_{v_0}, z_{v_1}, \dots, z_{v_T}\}$ ,

the registration network will decode a time-sequence of  $\{v_0, v_1, \dots, v_\tau\}$  accordingly.

Given a set of  $N$  observations,  $\hat{v}^n \triangleq \{\hat{v}_1^n, \dots, \hat{v}_\tau^n\}$ , which are numerical solutions to the geodesic shooting equation (Eq. (2)) with a given initial condition, we formulate our geodesic loss as the mean squared error (MSE) between the predicted time-dependent velocities and the corresponding velocities computed via the numerical solution to the EPDiff equation, i.e.

$$l(\theta_r, \theta_v) = \sum_{n=1}^N \|\mathcal{D}_{\theta_v}(\mathcal{G}_{\theta_r}(z_{v_0})) - \hat{v}^n\|_2^2 + \text{Reg}(\theta_r, \theta_v), \quad (5)$$

where  $\mathcal{G}_{\theta_r}(z_{v_0})$  denotes the geodesic mapping function that transforms the latent initial velocity  $z_{v_0}$  into a sequence of time-dependent latent velocities  $z_{v_1}, \dots, z_{v_\tau}$ . The decoder  $\mathcal{D}_{\theta_v}(\cdot)$  then maps this latent trajectory back to the image space, reconstructing the deformation process. The term  $\text{Reg}(\cdot)$  represents a regularization applied to the network parameters.

### 3.2 Joint Learning of GDN

The loss function of GDN integrates losses from both an unsupervised registration network and the geodesic neural operators informed by real numerical solutions. Given a set of pairwise images  $\{S^n, T^n\}_{n=1}^N$ , we are now ready to define the joint loss of GDN as

$$L(\theta_r, \theta_v) = \sum_{n=1}^N \lambda \|(S^n(\phi_\tau^n(\theta_v)) - T^n)\|_2^2 + \frac{1}{2} \xi (\mathcal{L}v_0^n(\theta_v), v_0^n(\theta_v)) + \eta l(\theta_r, \theta_v) \quad \text{s.t. Eq. (1)}, \quad (6)$$

where  $\lambda$ ,  $\xi$ , and  $\eta$  are positive weighting parameters to balance the image matching term and the geodesic loss defined in Eq. (5).

## 4. Experimental Evaluation

We demonstrate the effectiveness of our model, GDN, on both 2D synthetic data and 3D brain MRI scans. Our experimental design focuses on three main perspectives to evaluate the performance of the proposed GDN. We first quantitatively measure the difference between our predicted geodesic path of deformations and the real numerical solutions. We then show the capability of our model in promoting the generalizability and regularizability of the registration network. We finally report the quantitative results of the registration accuracy. Additionally, it is worth noting that the prediction of geodesics designed in GDN bypasses

the need for computing numerical solutions of EPDiff in previous learning-based LDDMM registration framework (Yang et al., 2017; Wang and Zhang, 2020; Hinkle et al., 2018), resulting in a reduced testing inference time.

### 4.1 Datasets

**2D synthetic data.** We demonstrate our model on three synthetic datasets: 2D synthetic circles with random radius, 2D hand-written digits MNIST (LeCun, 1998), and Google QuickDraw (Jongejan et al., 2016) - a collection of categorized drawings contributed by online players in a drawing game. For the QuickDraw dataset, we randomly choose 10000 images that include different classes of envelopes, moons, triangles, shirts, etc. All images were upsampled to the size of  $64 \times 64$ , and were pre-aligned with affine transformations within each class.

**3D brain MRI.** We include 828 T1-weighted 3D brain MRI scans from the Open Access Series of Imaging Studies (OASIS-3) dataset (Foteno et al., 2005). This dataset is part of the Learn2Reg challenge (Alsinan et al., 2022) and has been widely used in the literature as a standard benchmark (Yang et al., 2017; Wang and Zhang, 2020; Hinkle et al., 2018; Balakrishnan et al., 2019; Chen et al., 2022). The MRI scans were resampled to the dimension of  $128 \times 128 \times 128$ , with an isotropic resolution of  $1.25\text{mm}^3$ . All MRIs have undergone skull-stripping, intensity normalization, bias field correction, and affine alignment. Due to the difficulty of preserving the diffeomorphic property across individual subjects, particularly with large age variations, we carefully evaluate images from subjects aged 60 to 90. We split the dataset into 70% for training, 15% for validation, and 15% for testing. The testing volumes include manually delineated anatomical structures, such as White-Matter (WM), Cerebral-Cortex (CerebralC), Ventricle (Ven), Cerebellum-Cortex (CerebellumC), Thalamus (Tha), Putamen (Put), Caudate (Caud), Hippocampus (Hipp), and brain stem (Stem).

### 4.2 Experimental Design

**Geodesic learning evaluation.** We evaluate the effectiveness of our proposed GDN in learning geodesic deformations on 2D synthetic data. Specifically, we randomly select 12,000 image pairs from a combined dataset of MNIST (2,000 per digit ranging from 0 to 4) and binary circles (2,000 pairs). The dataset is partitioned into 70% for training, 15% for validation, and 15% for testing. We compare the predicted velocity fields,  $\{v_t\}$ , and their associated transformations,  $\{\phi_t\}$ , with two optimization-based algorithms - (i) the numerical solutions derived from the original geodesic shooting equation (Eq. (2)) and (ii) predictions from NeurEPDiff (Wu and Zhang, 2023). We also visualize

a sequence of deformed images along the predicted geodesic path.

**Evaluation on generalizability and regularizability.** To investigate the generalizability and regularizability of GDN in the context of registration, we first train the model on 2D circles and 2D MNIST data (digits 0-4 only) and test it on data from the same dataset but were not seen during training. We then evaluate GDN’s performance across different distributions of training and testing data. This includes testing the trained model on the remaining digits (digits 5-9) and a distinct dataset from Google QuickDraw. The final deformed images vs. target images are produced to examine the quality of image alignment. Meanwhile, we visualize the predicted deformations and their determinant of Jacobian maps to examine whether they are well regularized. Additionally, we report the percentage of voxels with a negative Jacobian determinant on the 2D dataset and the negative values observed in the determinant of Jacobian maps across all methods for real 3D brain MRI datasets. We compare the performance of GDN with the state-of-the-art deep learning-based diffeomorphic registration methods, notably LagoMorph (LM) (Hinkle et al., 2018), QuickSilver (QS) (Yang et al., 2017), VoxelMorph-diff (VM) (Dalca et al., 2019), and TransMorph (TM) (Chen et al., 2022). Please note that only the LM and QS methods are based on LDDMM. The other two baseline approaches are based on stationary velocity fields and do not necessarily generate geodesic paths unless the space is equipped with a flat affine connections (Lorenzi and Pennec, 2013). All methods are trained with their best performance reported.

**Final deformation evaluation.** We evaluate the quality of predicted deformations of GDN by quantitatively measuring its accuracy on real 3D brain MRIs. We perform registration-based segmentation and compare the resulting segmentation accuracy of our GDN with all baseline algorithms using two distinct quantification metrics. One commonly used metric is the dice similarity coefficient (DSC) (Dice, 1945) which evaluates volume overlap between the propagated/deformed segmentation  $A$  and the manual segmentation  $B$  for each structure. Such a dice score can be computed by  $DSC(A, B) = 2(|A \cap B|)/(|A| + |B|)$ , where  $\cap$  denotes an intersection of two regions.

**Statistical evaluation.** We performed paired t-tests (Student, 1908) to compare our proposed model GDN with other benchmarks and reported the corresponding p-values for all models on the DSC scores across different anatomical structures using real 3D brain MRI datasets. This analysis provides insight into whether GDN performs significantly better or differently compared to the baselines in terms of registration accuracy.

**Parameter setting.** The training of all our experiments is implemented on a server with AMD EPYC 7502 CPU of 126GB memory and Nvidia GTX 3090Ti GPUs. We train 2D and 3D networks using the Adam optimizer (Kingma and Ba, 2014) with weight decay as  $1e^{-4}$ . The 2D network was trained for 1000 epochs with a batch size of 100, while the 3D network was trained for 2000 epochs with a batch size of 6. We used a learning rate of  $5e^{-4}$  for our network GDN. We set the number of integration steps to  $\tau = 10$  for the 2D dataset and  $\tau = 5$  for the 3D dataset to balance accuracy and efficiency. The smoothing parameters for the operator  $K$  in Eq. (2) are set as  $\alpha = 1.0$ ,  $\beta = 0.5$  for the 2D dataset, and  $\alpha = 1.0$ ,  $\beta = 1.0$  for the 3D dataset. The positive weighting parameter for registration is  $\lambda = \frac{1}{0.03^2}$ , with the regularization weight set to  $\xi = 2.0$  for 2D and  $\xi = 0.15$  for 3D. The number of hidden layers in the GNO is configured as  $J = 4$  for 2D and  $J = 1$  for 3D. The backbone of the registration network is a U-Net architecture with a kernel size of 3. For our GDN, the channel configurations are set to  $[[8, 16, 16, 8], [8, 16, 16, 16, 8]]$ . All baseline models, including LM, QS, and VM-diff, use the same U-Net architecture with identical kernel size and channel configurations to ensure a fair comparison.

**Ablation Study.** We conduct an ablation study on 3D real brain MRIs to evaluate the influence of the number of integration steps ( $\tau$ ) in the geodesic learning module on both registration accuracy and regularity.

Please note that, our model GDN, when excluding the proposed GNO module and geodesic loss, is equivalent to the baseline LagoMorph. This serves as an ablation study to assess the contribution of the GNO and geodesic loss components.

**Numerical scheme in GNO learning.** The numerical solution of the geodesic shooting equation (Eq. (2)) that guides the learning of GNO is generated in real-time during the training process. In all experiments, we opt for a commonly employed Euler integrator to numerically solve this shooting equation (Yang et al., 2017; Wang and Zhang, 2020; Hinkle et al., 2018; Vialard et al., 2012b; Beg et al., 2005). Other advanced integrators, such as Runge-Kutta, can be easily applied.

### 4.3 Experimental Results

In Fig. 2, we present the visualizations of the predicted transformations from GDN alongside numerical solution obtained from two optimization-based algorithms, the original geodesic shooting equation and NeurEPDiff. The deformation processes generated by both GDN and numerical solution show a nice degree of similarity, while GDN slightly outperforms NeurEPDiff.

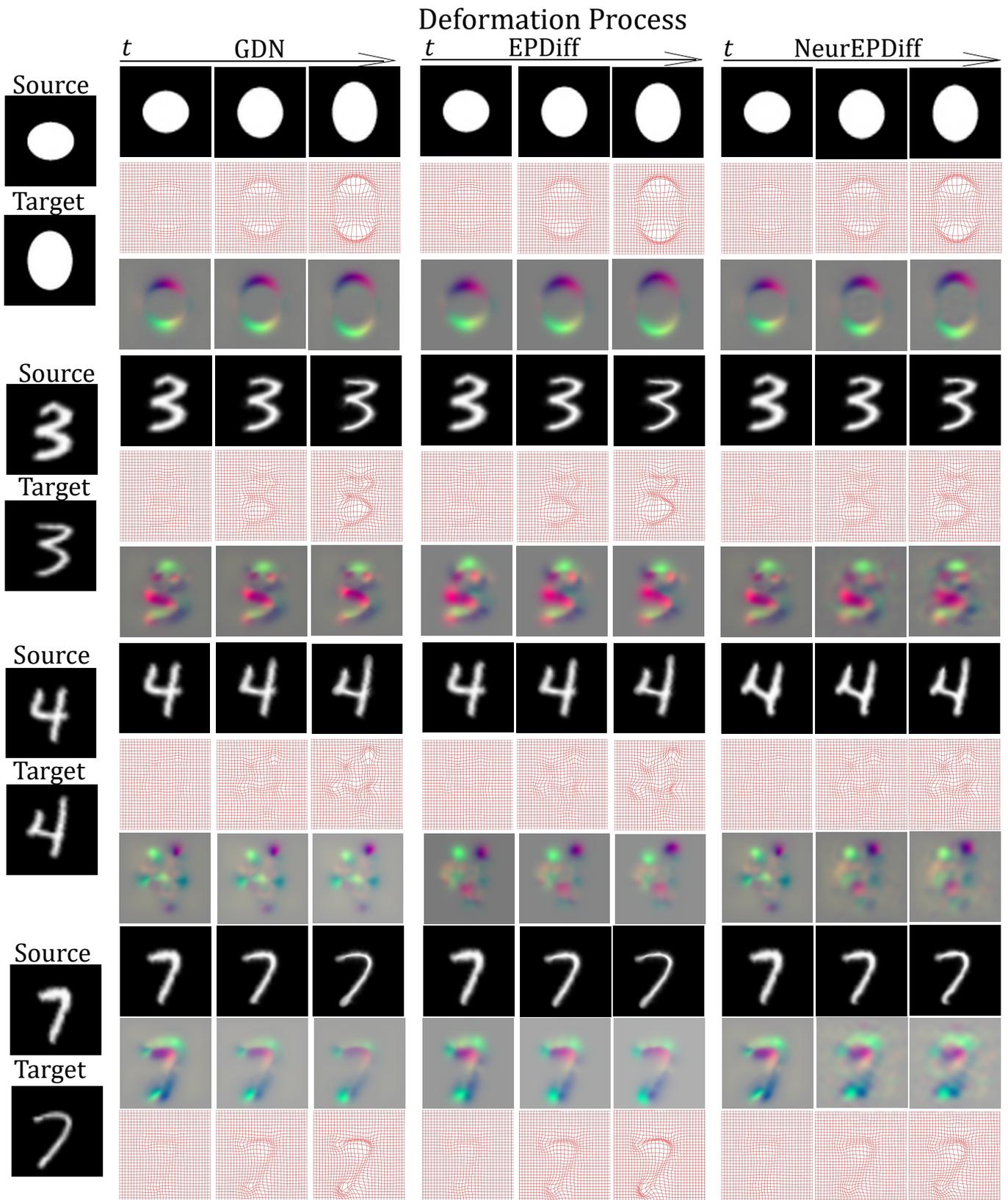


Figure 2: Visualization of predicted geodesics from GDN vs. real numerical solutions (Zhang and Fletcher, 2015) and predictions from NeurEPDiff (Wu and Zhang, 2023). Left to right: source and target images, predicted geodesic deformations along time  $t$ . Top to bottom: deformed images, transformation fields, and velocity fields.

Fig. 3 compares the discrepancy between our method and the numerical solution to the EPDiff equation by using NeurEPDiff as a baseline for reference. Specifically, we quantify the mean squared error (MSE) between the

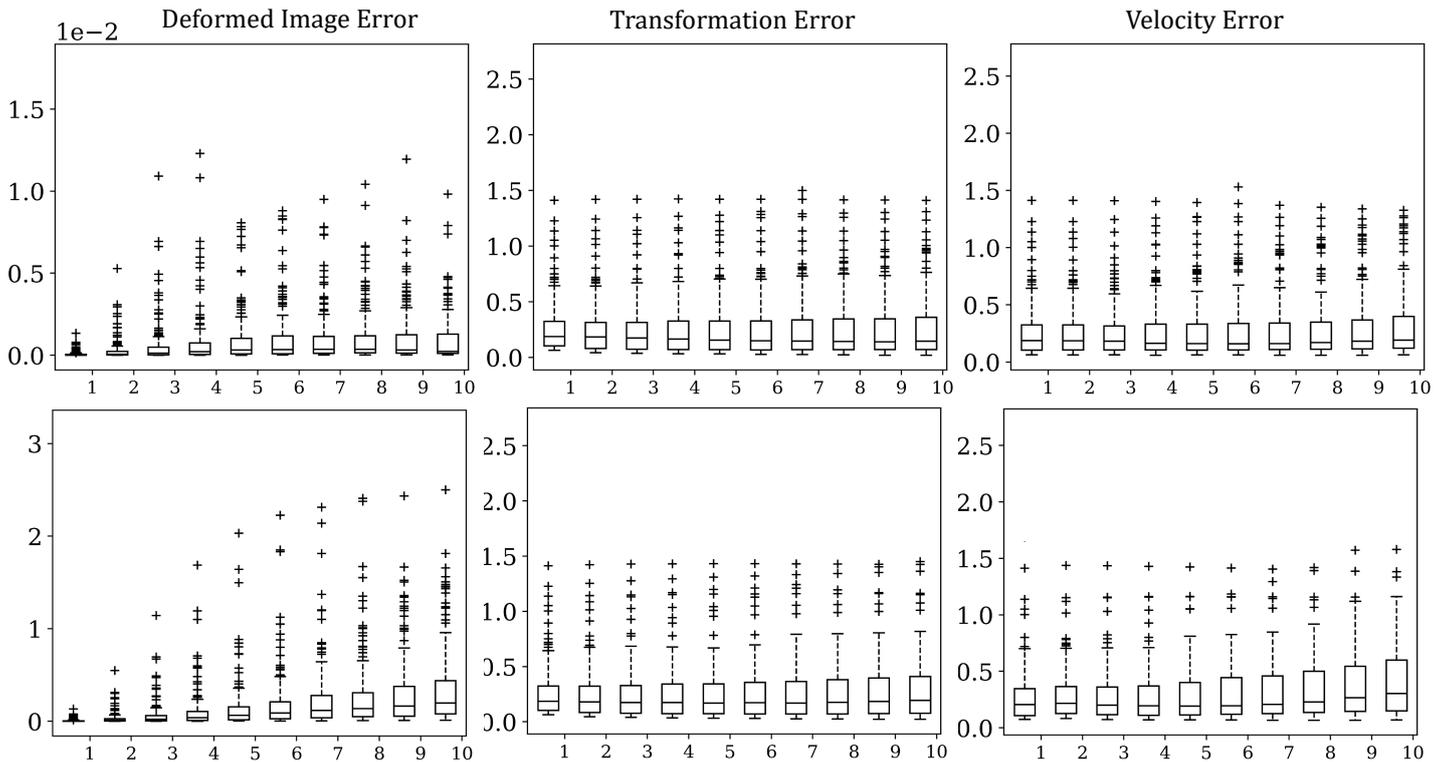


Figure 3: Top to bottom: Comparison of predicted geodesics by GDN/NeurEPDiff vs. numerical integration of EPDiff. Left to right: MSE between deformed images, transformations, and velocities along the geodesic path.

predicted time-varying velocity fields and the geodesic velocities obtained through numerical integration of EPDiff. Additionally, we assess the similarity of the resulting transformations and deformed images along the predicted and ground-truth geodesic paths. This result shows that our method, GDN, achieves slightly better performance than NeurEPDiff across these metrics.

Fig. 4 visualizes examples of predicted transformation fields and their associated determinant of Jacobian (DetJac) maps on in-distribution (ID) testing data from all methods. Different values of DetJac indicate different patterns of volume changes. For example, a DetJac value of 1 indicates no volume change, while  $\text{DetJac} < 1$  reflects volume shrinkage and  $\text{DetJac} > 1$  implies volume expansion. A DetJac value smaller than zero suggests an artifact or singularity in the transformation field, i.e., a failure to regularize the smoothness of deformation fields when the effect of folding and crossing grids occurs. While the value of DetJac from all methods do not suggest violations of regularization, our work GDN exhibits smoother transformation grids, indicating more effective regularization. In particular, the estimated transformations from GDN are well regularized to concentrated primarily on the shape of circles or digits. Deformation changes on background areas (particularly with no intensity changes between images) have little-to-zero values.

Fig. 5 displays a comparison of predicted transformation fields generated from out-of-distribution testing (OOD) data. The results indicate that GDN consistently produces smoother transformation grids, ensuring well-aligned source-to-target images. These consistent findings suggest that our introduced geodesic learning approach in image registration contributes to an enhanced network generalizability and regularizability compared to the other baseline methods.

Fig. 6 presents visual comparisons of deformed images, registration-based segmentation maps overlaid on 3D brain MRIs, predicted deformation fields, and determinants of Jacobian maps for all methods, shown in both sagittal and coronal views. Our method demonstrates a similar smoothing effect on transformation grids with LM and QS, while more clearly highlighting biologically plausible growth patterns, such as ventricular enlargement (highlighted in red in the determinants of Jacobian maps). These patterns are more interpretable than those produced by the baseline methods, VM and TM.

Fig. 7 displays a quantitative comparison between the manually labeled vs. propagated segmentation labels deformed by deformations fields predicted from all methods. The top panel reports the statistics of dice scores (the higher the better) over eight brain structures from hundreds of testing registration pairs. The bottom panel summarizes the corresponding means and standard deviations. It shows

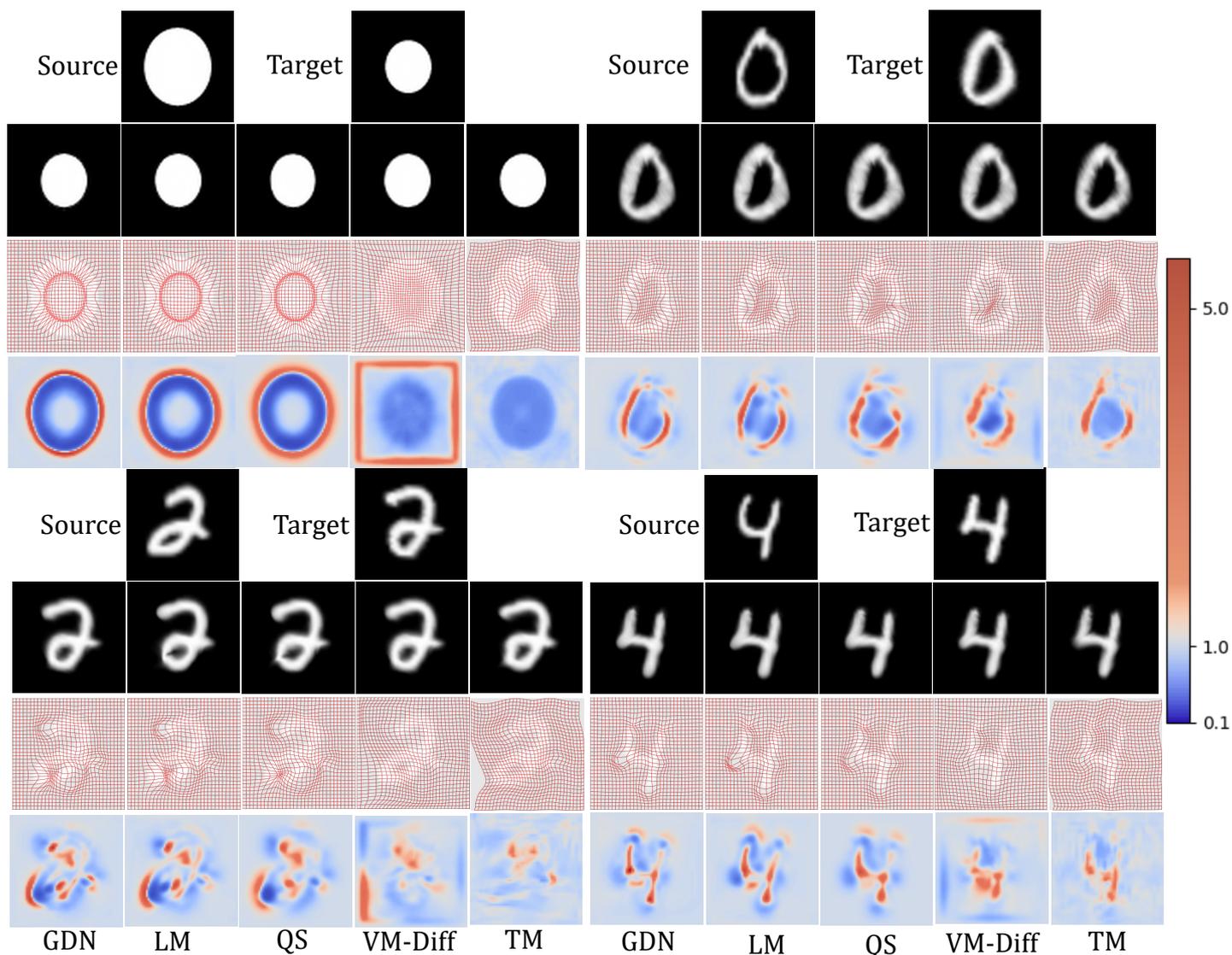


Figure 4: A comparison of predicted transformation grids and their associated determinant of Jacobian maps on in-distribution (ID) testing data from our method GDN vs. baselines. From top to bottom: source and target image pairs, resulting deformed images, predicted transformation grids, and determinant of Jacobian maps.

that our method GDN is able to produce superior-quality final deformation fields in comparison with other baselines. Overall, our GDN achieves notably higher Dice on average. Meanwhile, an additional benefit of GDN is to forgo the computation of the geodesic shooting equation inherent in conventional LDDMM-based registration networks. This step can substantially increase computational cost in the testing phase, particularly when high-dimensional images. In this experiment, GDN demonstrates its capability to predict geodesics in 0.026s, outperforming the other two baselines in the context of LDDMM - LM and QS (0.036 seconds). It's important to note that the other two baselines - VM-diff (Balakrishnan et al., 2019) and TM (Chen et al., 2022) models employ stationary velocity fields to parameterize the deformation fields, treating the velocities as constant over time.

We also report the corresponding p-values from pairwise statistical tests comparing our proposed model, GDN, with four baseline methods across the anatomical brain structures based on Dice scores. The p-values for GDN vs. LM ( $< 0.001$ ), QS ( $< 0.001$ ), VM ( $< 0.001$ ), and TM ( $< 0.001$ ) are all well below the significance threshold of 0.05, indicating that GDN performs significantly differently compared with baselines.

Fig. 8 presents a quantitative comparison of negative values in the determinant of Jacobian (DetJac) of the transformations across three different datasets — in-distribution (ID) testing data, out-of-distribution (OOD) testing data, and real 3D brain MRIs. The results indicate that for both the ID testing data and the 3D brain MRIs, there are barely violations of regularization for all methods. However, it is worth noting that our method, GDN, demonstrates

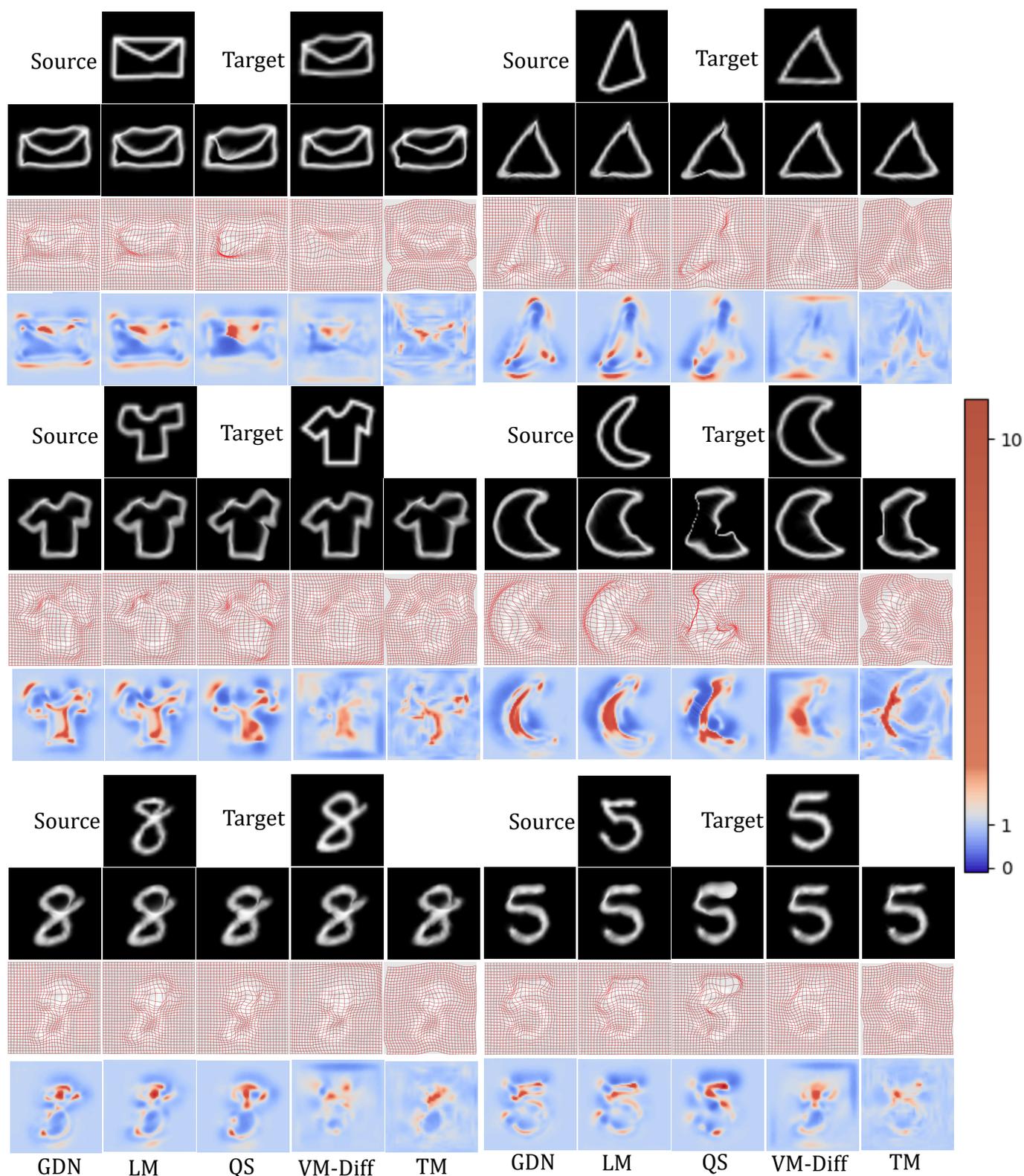


Figure 5: A comparison of predicted transformation grids and their associated determinant of Jacobian maps on out-of-distribution (OOD) testing data from our method GDN vs. baselines. Top to bottom: source and target images, deformed source images, predicted deformation fields, and determinant of Jacobian maps.

advantages on the OOD testing data, with zero negative values reported. However, when tested on data outside the training distribution, the percentage of negative Jacobian determinants begins to diverge across models: GDN (0.005%), LM (0.132%), QS (0.255%), VM-Diff (0.038%), and TM (0.011%). These findings indicate the potential of GDN to achieve improved generalizability and regularization compared to other baselines.

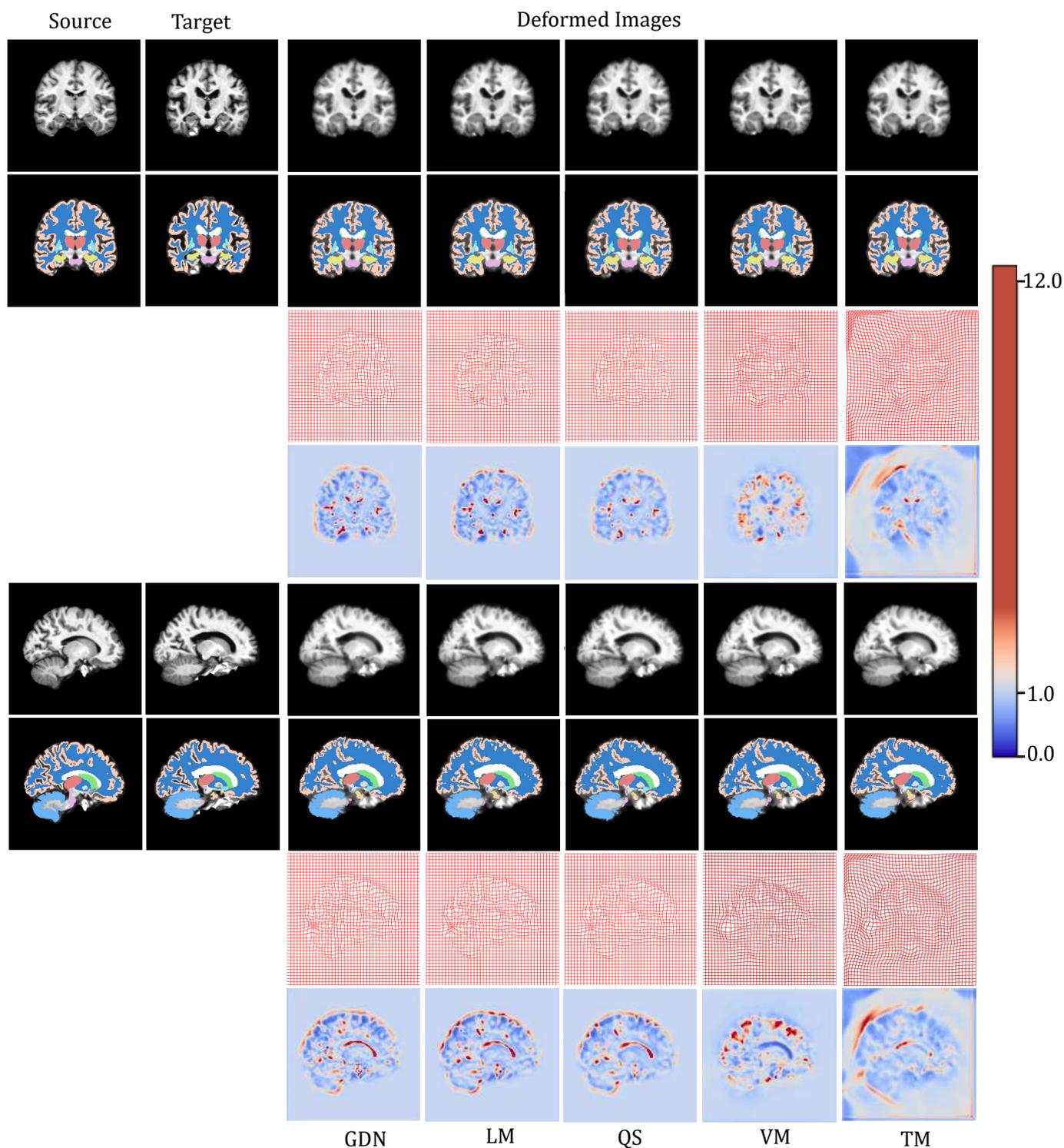


Figure 6: Top to bottom: examples of 3D brain MRIs (demonstrated in coronal and sagittal views), registration-based segmentation maps overlaid on brain MRIs, predicted deformation fields, and determinants of Jacobian maps. Left to right: source and target images; deformed source images (segmentation labels) across all methods.

Tab. 1 summarizes the model parameters, training time per epoch, and inference time for GDN versus the baselines. Our model, GDN, has a larger number of parameters than LM and QS (under the same LDDMM setting) due to the added geodesic learning module. In addition, training GDN requires solving the EPDiff equation numerically to

supervise the geodesic loss. These two factors contribute to GDN's longer training time per epoch compared to LM and QS. However, a key advantage of GDN lies in inference: once trained, GDN does not require numerical integration during testing. As a result, GDN achieves faster inference time (0.030s) than LM and QS (0.043s), and we expect

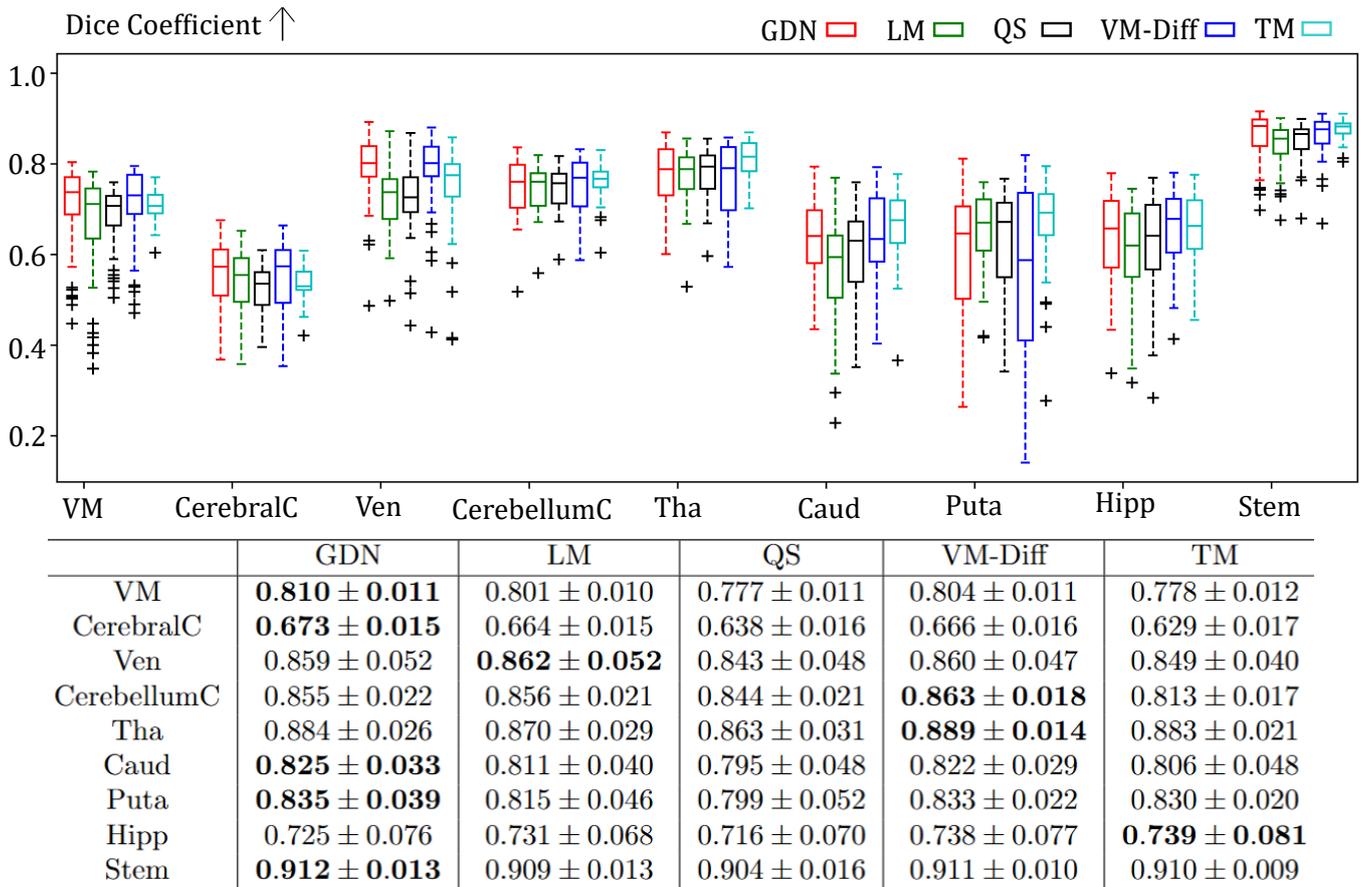


Figure 7: Top: A comparison of dice score by propagating the deformation field to the segmentation labels on nine brain structures White-Matter (WM), Cerebral-Cortex (CerebralC), Ventricle (Ven), Cerebellum-Cortex (CerebellumC), Thalamus (Tha), Putamen (Putamen), Caudate (Caud), Hippocampus (Hipp), and brain stem (Stem).

	LDDMM-Based			SVF-Based	
	GDN	LM	QS	VM-Diff	TM
Number of Parameters	79.3K	56.5K	56.5K	56.5K	46.5M
Training (per epoch)	61.5s	14.61s	20.5s	20.78s	33.39s
Testing	0.030s	0.043s	0.043s	0.017s	0.020s

Table 1: A summary of model parameters, training time per epoch, and testing time on 3D brain MRI data.

this advantage to become more pronounced with higher-dimensional image data. When compared to VM-diff and TM under the stationary velocity field setting, GDN does not show a significant advantage in training or inference time, primarily due to its larger model size relative to VM-diff. The model TM, while larger in size, benefits from parallel computation and hence is more time-efficient.

Fig. 9 presents a quantitative comparison of the average

dice scores on 3D brain MRIs with different integration steps of our proposed module GNO. The results show that our method achieves stable dice accuracy based on the deformed template segmentation labels as the number of integration steps increases from five to ten.

## 5. Conclusions

This paper presents a novel deep network, named as GDN, that *for the first time learns geodesics of deformation spaces in learning-based image registration methods*. In contrast to current approaches that solely fit the registration loss, we develop a geodesic deformation subnetwork guided by the fundamental physical processes of deforming images from one to another. To achieve this, we treat the geodesics as unknown mapping functions directly learned from the data via a carefully developed neural operator. A composition of integral operators and smooth activation functions is then formulated to effectively approximate such mappings. Our proposed GDN jointly optimizes a newly defined geodesic loss function with an alternating optimization scheme. Ex-

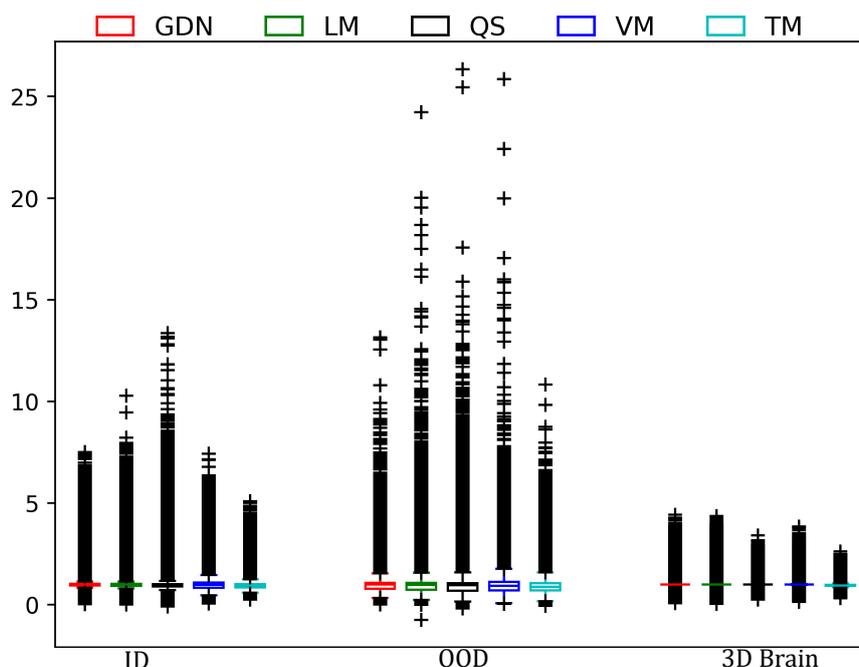


Figure 8: A report on the negative values in determinant of Jacobian of the transformations on three different datasets: in-distribution (ID) testing data, out-of-distribution (OOD) testing data, and real 3D brain MRIs.

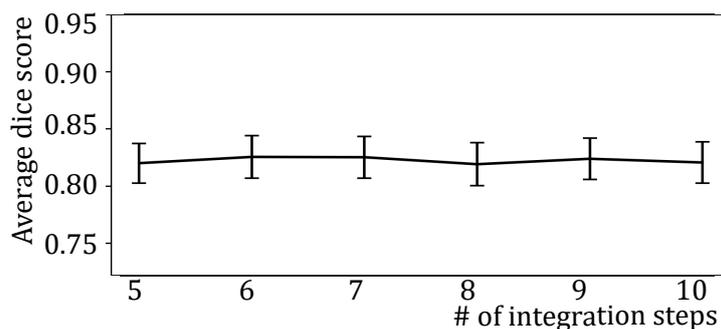


Figure 9: A comparison of average dice score on 3D brain MRIs with a varying number of integration steps.

perimental results on both 2D synthetic data and 3D brain MRI scans show that our model gains an improved generalizability with more effective regularizations on the final deformation fields.

Our work GDN is an initial step towards quantifying and analyzing geometric deformations from images within deep neural networks. The proposed geodesic learning lays a foundation for performing further quantitative and interpretable analysis in the latent spaces of deformation fields. Our potential future works may focus on learning geodesic regression and population-based variability in deep nets.

## Acknowledgments

This work was supported by NSF CAREER Grant 2239977.

## Ethical Standards

The work follows appropriate ethical standards in conducting research and writing the manuscript, following all applicable laws and regulations regarding treatment of animals or human subjects.

## Conflicts of Interest

We declare we do not have conflicts of interest.

## Data availability

The Brain MRIs used in training and testing in this paper are from the Open Access Series of Imaging Studies (OASIS-3) dataset (LaMontagne et al., 2019), which are readily accessible and user-friendly. Readers interested in evaluating

the accuracy of the method can access and utilize the publicly available datasets. The code for this study is also publicly available at <https://github.com/nellie689/GDN>, with an archived version on Zenodo at <https://zenodo.org/records/17733235>.

## References

- Ahmed Alsinan, Richard Fan, Geoffrey Sonn, and Mirabela Rusu. The learn2reg 2021 miccai grand challenge (pimed team). *Biomedical Image Registration, Domain Generalisation and Out-of-Distribution Analysis: MICCAI 2021 Challenges: MIDOG 2021, MOOD 2021, and Learn2Reg 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27–October 1, 2021, Proceedings*, 13166:168, 2022.
- Vladimir Arnold. Sur la géométrie différentielle des groupes de lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. In *Annales de l'institut Fourier*, volume 16, pages 319–361, 1966.
- Brian B Avants, Charles L Epstein, Murray Grossman, and James C Gee. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis*, 12(1):26–41, 2008.
- Guha Balakrishnan, Amy Zhao, Mert R Sabuncu, John Guttag, and Adrian V Dalca. Voxelmorph: a learning framework for deformable medical image registration. *IEEE transactions on medical imaging*, 2019.
- Han Bao, Hui Ren, Zhiling Zhou, Xiang Li, Ning Guo, and Quanzheng Li. 3d regional shape analysis of left ventricle using mr images: Abnormal myocardium detection and classification. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 789–792. IEEE, 2019.
- Chris Beekman, Suzanne van Beek, Jikke Stam, Jan-Jakob Sonke, and Peter Remeijer. Improving predictive ctv segmentation on ct and cbct for cervical cancer by diffeomorphic registration of a prior. *Medical Physics*, 49(3):1701–1711, 2022.
- Mirza Faisal Beg, Michael I Miller, Alain Trouvé, and Laurent Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International journal of computer vision*, 61(2):139–157, 2005.
- Kaushik Bhattacharya, Bamdad Hosseini, Nikola B Kovachki, and Andrew M Stuart. Model reduction and neural networks for parametric pdes. *The SMAI journal of computational mathematics*, 7:121–157, 2021.
- Junyu Chen, Eric C Frey, Yufan He, William P Segars, Ye Li, and Yong Du. Transmorph: Transformer for unsupervised medical image registration. *Medical image analysis*, 82:102615, 2022.
- Gary E Christensen, Richard D Rabbitt, and Michael I Miller. A deformable neuroanatomy textbook based on viscous fluid mechanics. In *27th Ann. Conf. on Inf. Sciences and Systems*, pages 211–216, 1993.
- P Constantin. *Navier-Stokes Equations*. University of Chicago Press, 1988.
- Adrian V Dalca, Guha Balakrishnan, John Guttag, and Mert R Sabuncu. Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces. *Medical image analysis*, 57:226–236, 2019.
- Lee R Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.
- Anthony F Fotenos, AZ Snyder, LE Girton, JC Morris, and RL Buckner. Normative estimates of cross-sectional and longitudinal brain volume decline in aging and ad. *Neurology*, 64(6):1032–1039, 2005.
- Ulf Grenander and Michael I Miller. Computational anatomy: An emerging discipline. *Quarterly of applied mathematics*, 56(4):617–694, 1998.
- Dan Hendrycks and Kevin Gimpel. Gaussian error linear units (gelus). *arXiv preprint arXiv:1606.08415*, 2016.
- J. Hinkle, D. Womble, and H.J.d Yoon. Diffeomorphic autoencoders for LDDMM atlas building. In *Medical Imaging with Deep Learning*, 2018.
- Sungmin Hong, James Fishbaugh, Jason J Wolff, Martin A Styner, Guido Gerig, and IBIS Network. Hierarchical multi-geodesic model for longitudinal analysis of temporal trajectories of anatomical shape and covariates. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part IV 22*, pages 57–65. Springer, 2019.
- Yi Hong, Polina Golland, and Miaomiao Zhang. Fast geodesic regression for population-based image analysis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 317–325. Springer, 2017.
- Jonas Jongejan, Henry Rowley, Takashi Kawashima, Jongmin Kim, and Nick Fox-Gieg. The quick, draw!-ai experiment. *Mount View, CA, accessed Feb*, 17(2018):4, 2016.

- Boah Kim, Inhwa Han, and Jong Chul Ye. Diffusemorph: Unsupervised deformable image registration using diffusion model. In *Computer Vision—ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXXI*, pages 347–364. Springer, 2022.
- Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces. *arXiv preprint arXiv:2108.08481*, 2021.
- Pamela J LaMontagne, Tammie LS Benzinger, John C Morris, Sarah Keefe, Russ Hornbeck, Chengjie Xiong, Elizabeth Grant, Jason Hassenstab, Krista Moulder, Andrei G Vlassenko, et al. Oasis-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and alzheimer disease. *medrxiv*, pages 2019–12, 2019.
- Yann LeCun. The mnist database of handwritten digits. <http://yann.lecun.com/exdb/mnist/>, 1998.
- Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*, 2020a.
- Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Graph kernel network for partial differential equations. *arXiv preprint arXiv:2003.03485*, 2020b.
- Zongyi Li, Daniel Zhengyu Huang, Burigede Liu, and Anima Anandkumar. Fourier neural operator with learned deformations for pdes on general geometries. *Journal of Machine Learning Research*, 24(388):1–26, 2023.
- Marco Lorenzi and Xavier Pennec. Geodesics, parallel transport & one-parameter subgroups for diffeomorphic image registration. *International journal of computer vision*, 105(2):111–127, 2013.
- Marco Lorenzi, Nicholas Ayache, Giovanni Frisoni, Xavier Pennec, et al. 4d registration of serial brain’s mr images: a robust measure of changes applied to alzheimer’s disease. In *Spatio Temporal Image Analysis Workshop (STIA), MICCAI*, volume 1. Citeseer, 2010.
- Lu Lu, Pengzhan Jin, and George Em Karniadakis. Deeponet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators. *arXiv preprint arXiv:1910.03193*, 2019.
- Michael I Miller. Computational anatomy: shape, growth, and atrophy comparison via diffeomorphisms. *NeuroImage*, 23:S19–S33, 2004.
- Michael I Miller, Alain Trounev, and L. Younes. Geodesic shooting for computational anatomy. *Journal of Mathematical Imaging and Vision*, 24(2):209–228, 2006.
- Marc Niethammer, Yang Huang, and François-Xavier Vialard. Geodesic regression for image time-series. In *International conference on medical image computing and computer-assisted intervention*, pages 655–662. Springer, 2011.
- Marc Niethammer, Roland Kwitt, and Francois-Xavier Vialard. Metric learning for image registration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8463–8472, 2019.
- Anqi Qiu, Laurent Younes, Michael I Miller, and John G Csernansky. Parallel transport in diffeomorphisms distinguishes the time-dependent pattern of hippocampal surface deformation due to healthy aging and the dementia of the alzheimer’s type. *NeuroImage*, 40(1):68–76, 2008.
- Anqi Qiu, Laurent Younes, and Michael I Miller. Principal component based diffeomorphic surface mapping. *Medical Imaging, IEEE Transactions on*, 31(2):302–311, 2012.
- Daniel Rueckert, Alejandro F Frangi, and Julia A Schnabel. Automatic construction of 3-d statistical deformation models of the brain using nonrigid registration. *IEEE transactions on medical imaging*, 22(8):1014–1025, 2003.
- Nikhil Singh, Jacob Hinkle, Sarang Joshi, and P Thomas Fletcher. A hierarchical geodesic model for diffeomorphic longitudinal shape analysis. In *International Conference on Information Processing in Medical Imaging*, pages 560–571. Springer, 2013.
- Student. The probable error of a mean. *Biometrika*, pages 1–25, 1908.
- Shanlin Sun, Kun Han, Deying Kong, Hao Tang, Xiangyi Yan, and Xiaohui Xie. Topology-preserving shape reconstruction and registration via neural diffeomorphic flow. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20845–20855, 2022.

- Shanlin Sun, Kun Han, Chenyu You, Hao Tang, Deying Kong, Junayed Naushad, Xiangyi Yan, Haoyu Ma, Pooya Khosravi, James S Duncan, et al. Medical image registration via neural fields. *Medical Image Analysis*, 97: 103249, 2024.
- François X Vialard, Laurent Risser, Daniel Rueckert, and Colin J Cotter. Diffeomorphic 3d image registration via geodesic shooting using an efficient adjoint calculation. *International Journal of Computer Vision*, 97(2):229–241, 2012a.
- François-Xavier Vialard, Laurent Risser, Daniel Rueckert, and Colin J Cotter. Diffeomorphic 3d image registration via geodesic shooting using an efficient adjoint calculation. *International Journal of Computer Vision*, 97:229–241, 2012b.
- Jian Wang and Miaomiao Zhang. Deepflash: An efficient network for learning-based medical image registration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.
- Jian Wang and Miaomiao Zhang. Geo-sic: Learning deformable geometric shapes in deep image classifiers. *The Conference on Neural Information Processing Systems*, 2022.
- Lei Wang, Faisal Beg, Tilak Ratnanather, Can Ceritoglu, Laurent Younes, John C Morris, John G Csernansky, and Michael I Miller. Large deformation diffeomorphism and momentum based hippocampal shape discrimination in dementia of the alzheimer type. *IEEE transactions on medical imaging*, 26(4):462–470, 2007.
- WM Wells, P Viola, H Atsumi, S Nakajima, and R Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical image analysis*, 1996.
- Nian Wu and Miaomiao Zhang. Neurepdiff: Neural operators to predict geodesics in deformation spaces. In *International Conference on Information Processing in Medical Imaging*, pages 588–600. Springer, 2023.
- Yifan Wu, Tom Z Jiahao, Jiancong Wang, Paul A Yushkevich, M Ani Hsieh, and James C Gee. Nodoe: A neural ordinary differential equation based optimization framework for deformable image registration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 20804–20813, 2022.
- Xiao Yang, Roland Kwitt, Martin Styner, and Marc Niethammer. Quicksilver: Fast predictive image registration—a deep learning approach. *NeuroImage*, 158:378–396, 2017.
- Laurent Younes, Felipe Arrate, and Michael I Miller. Evolutions equations in computational anatomy. *NeuroImage*, 45(1):S40–S50, 2009.
- Miaomiao Zhang and P Thomas Fletcher. Finite-dimensional lie algebras for fast diffeomorphic image registration. In *International Conference on Information Processing in Medical Imaging*, pages 249–260. Springer, 2015.