NODEO: A Neural Ordinary Differential Equation Based Optimization Framework for Deformable Image Registration

Yifan Wu*, Tom Z. Jiahao*, Jiancong Wang, Paul A. YushKevich, M. Ani Hsieh, James C. Gee

University of Pennsylvania, Philadelphia, PA, USA

PENN IMAGE COMPUTING & SCIENCE LABORATORY

CVPR JUNE 19-24 2022 NEW ORLEANS • LOUISIANA

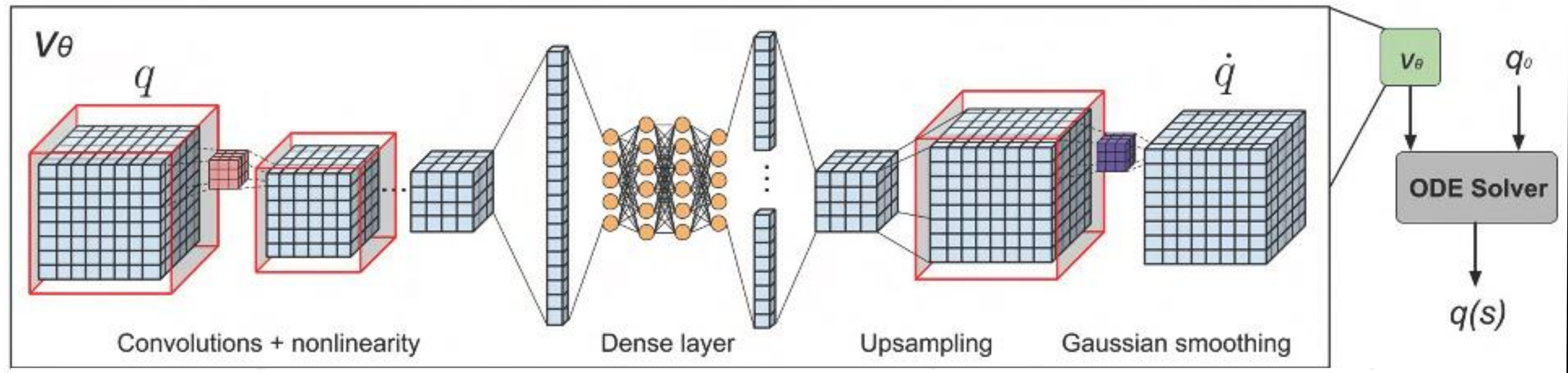
Penn Engineering | GRASP Laboratory

General Robotics, Automation, Sensing & Perception Lab

Overview

- Deformable image registration (DIR), aiming to find spatial correspondence between images, is one of the most critical problems in the domain of medical image analysis.
- In this paper, we present a novel and generic diffeomorphic image registration framework that utilizes neural ordinary differential equations (NODEs). We model each voxel as a moving particle and consider the set of all voxels in a 3D image as a high-dimensional dynamical system whose trajectory determines the targeted deformation field.
- Our method leverages deep neural networks for their expressive power in modeling dynamical systems, and simultaneously optimizes for a dynamical system between the image pairs and the corresponding transformation.
- Our formulation allows various constraints to be imposed along the transformation to maintain desired regularities. Our experiment results show that our method outperforms the benchmarks under various metrics.
- Additionally, we demonstrate the feasibility to expand our framework to register multiple image sets using a unified form of transformation, which could possibly serve a wider range of applications.

Method



We denote the location of all voxels or the voxel cloud in an image as the the ordered set q. The goal of our work is to find a transformation ψ which maps the domain of the voxel cloud onto itself, such that the transformed moving image is similar to the fixed image with desired constraints.

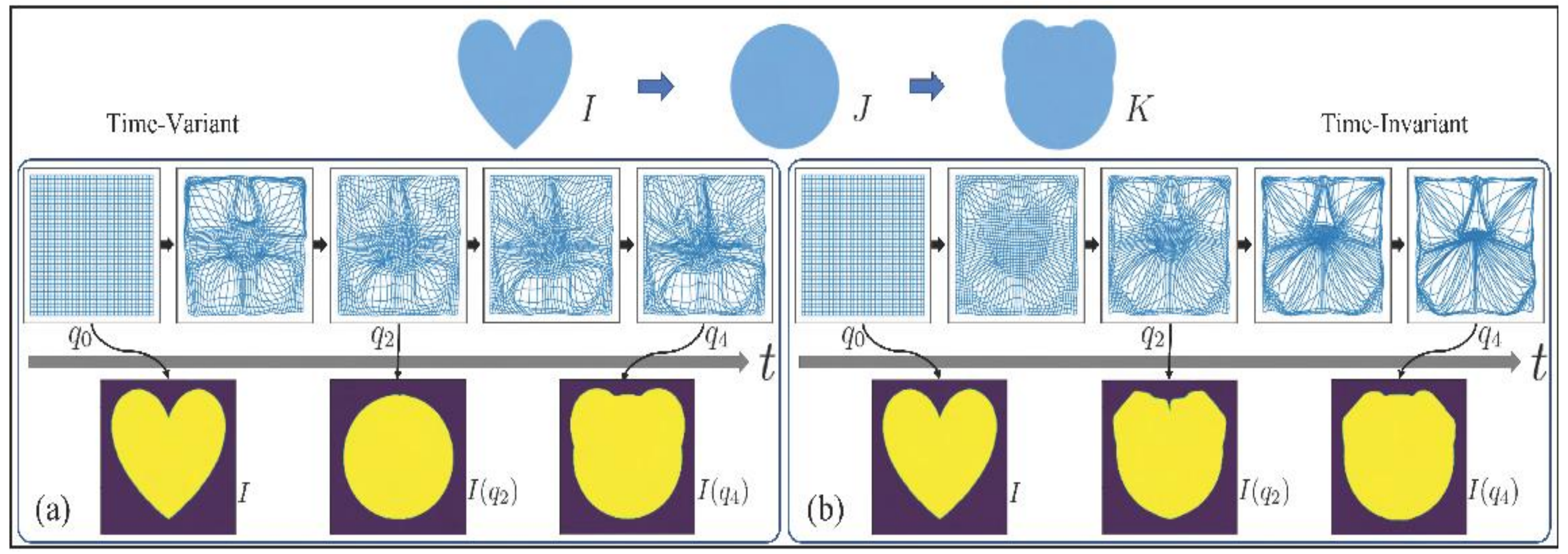
$$\frac{dq}{dt} = \mathcal{K}\mathbf{v}_\theta(q(t), t), \quad s.t.\ q(0) = q_0,$$

where $\mathbf{v}_\theta(\cdot)$, as parametrized by θ, is the vector field describing the dynamics of the voxel cloud, q_0 is the initial condition at $t = 0$. We employ Gaussian kernels (for $\Omega \subseteq \mathbb{R}^3$, we use 3D Gaussian kernels), denoted by $\mathcal{K}$, as a filtering operator to enforce spatial smoothness in Ω.

The trajectory of q is generated by integrating the ODE in equation above with the initial condition q_0. The resulting voxel cloud at $t = s$ denotes the transformation $\psi(q_0)$ given by:

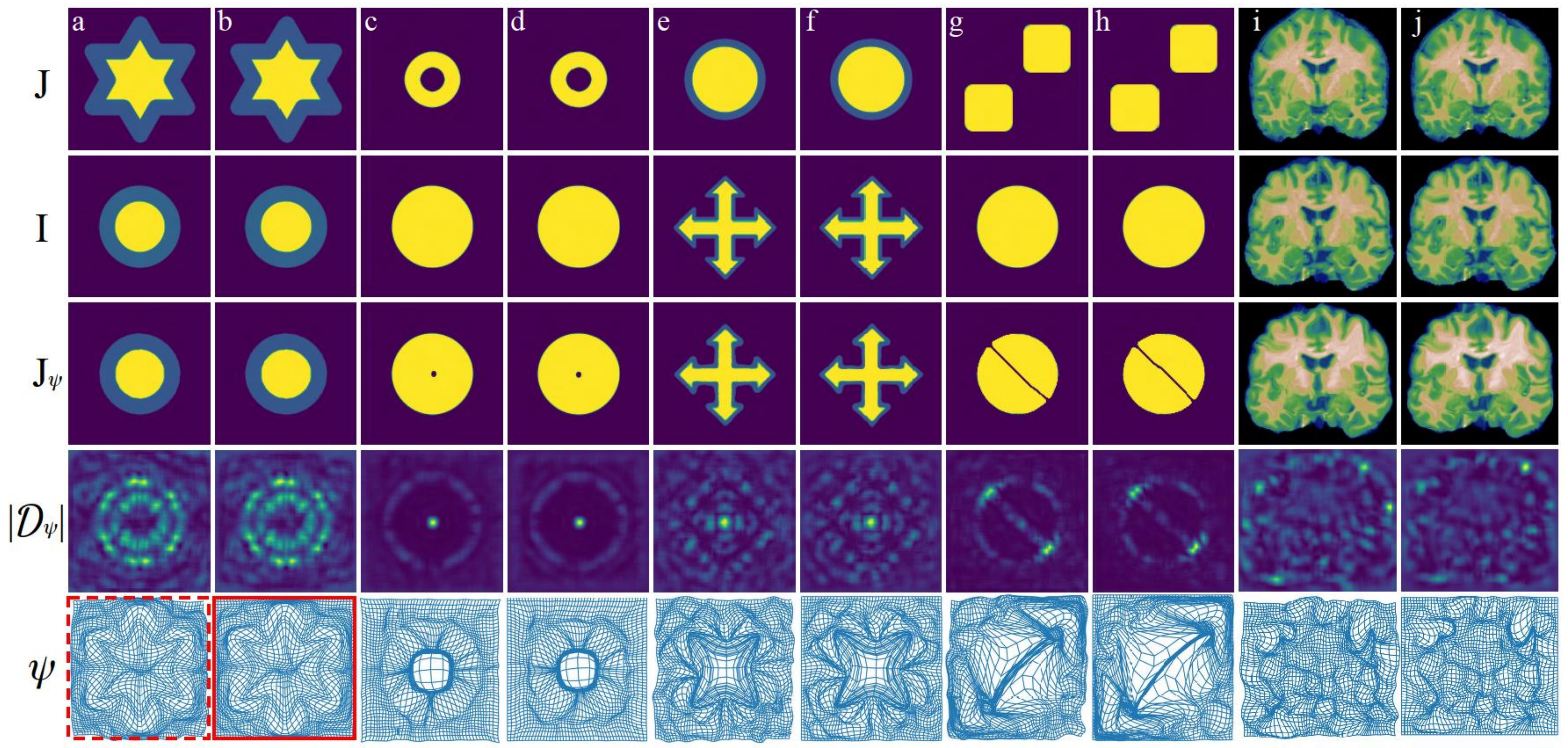
$$\psi(q_0) = q(s) = q_0 + \int_0^s \mathcal{K}\mathbf{v}_\theta(q(t), t)dt.$$

Time-Variant Modeling



Discovering transformation on multiple images. The target images I, J, K are shown in the top row, and the model is tasked to identify a path of transformation from I to K via J. The pictured registration results are from using (a) a time-varying system with explicit time embedding, and (b) a time-invariant system where the transformation does not depend on time. The transformations $\psi(q_t)$ are plotted on top of the warped images during the registration.

Illustrative Examples in 2D pair images



J: moving images
I: fixed images
J_ψ: warped moving images
$|\mathcal{D}_\psi|$: Jacobian determinants of ψ
ψ: deformation fields

We show our framework provide properties of:
topology preserving
(c, g present transformation between different topologies);
great expressive power under diffeomorphism
(a, c show smooth approximation of non-smooth manifolds);
desired constraints can be imposed
(b, d, f, h, j show registration with boundary conditions).

Results

OASIS dataset	Avg. Dice (28) ↑	$\mathcal{D}_\psi(\mathrm{x}) \le 0\ (r^{\mathcal{D}})$ ↓	$\mathcal{D}_\psi(\mathrm{x}) \le 0\ (s^{\mathcal{D}})$ ↓
SYMNet [25]	0.743 ± 0.113	0.026%	-
SyN [1]	0.729 ± 0.109	0.026%	0.005
NiftyReg [2]	0.775 ± 0.087	0.102%	1395.988
Log-Demons [3]	0.764 ± 0.098	0.121%	84.904
NODEO (ours $\lambda_1 = 2.5$)	0.778 ± 0.026	0.030%	34.183
NODEO (ours $\lambda_1 = 2$)	$\mathbf{0.779 \pm 0.026}$	0.030%	61.105
CANDI dataset	Avg. Dice (28) ↑	$\mathcal{D}_\psi(\mathrm{x}) \le 0\ (r^{\mathcal{D}})$ ↓	$\mathcal{D}_\psi(\mathrm{x}) \le 0\ (s^{\mathcal{D}})$ ↓
SYMNet [25]	0.778 ± 0.091	1.4×10^{-4}%	1.043
SyN [1]	0.739 ± 0.102	0.018%	0.012
NiftyReg [2]	0.775 ± 0.088	0.101%	1395.987
Log-Demons [3]	0.786 ± 0.094	0.071	49.274
NODEO (ours $\lambda_1 = 2.5$)	0.801 ± 0.011	7.5×10^{-8}%	1.574
NODEO (ours $\lambda_1 = 2$)	$\mathbf{0.802 \pm 0.011}$	1.8×10^{-7}%	4.341
CANDI dataset	Avg. Dice (32) ↑	$\mathcal{D}_\psi(\mathrm{x}) \le 0\ (r^{\mathcal{D}})$ ↓	$\mathcal{D}_\psi(\mathrm{x}) \le 0\ (s^{\mathcal{D}})$ ↓
SYMNet [25]	0.736 ± 0.015	1.4×10^{-4}%	1.043
SyN [1]	0.713 ± 0.177	0.018%	0.012
NiftyReg [2]	0.748 ± 0.160	0.101%	1395.987
Log-Demons [3]	0.744 ± 0.160	0.071	49.274
NODEO (ours $\lambda_1 = 2.5$)	$\mathbf{0.760 \pm 0.011}$	7.5×10^{-8}%	1.574
NODEO (ours $\lambda_1 = 2$)	$\mathbf{0.760 \pm 0.011}$	1.8×10^{-7}%	4.341

Evaluation metric: Dice Similarity Coefficient, Negative Jacobian Determinant.
Two brain MRI datasets: OASIS and CANDI.
Image Size: $160 \times 192 \times 144$. Total number of model parameters: around 3/4 of the number of voxels in the image.
Runtime: around 80s for one pair. GPU memory: 3863MB.