Descriptor-based Structural Similarity and Neural ODEs for Multimodal Diffeomorphic Registration

Salvador Rodriguez-Sanz[1], Mónica Hernández[1]

[1] Affiliation: T64_23R COSMOS, Computer Science for Complex Systems Modelling
Instituto de Investigación en Ingeniería de Aragón (I3A)
Universidad de Zaragoza, Mariano Esquillor s/n, 50018, Zaragoza, Spain.
Tel. +34-976762707, e-mail: srodsanz@unizar.es

Abstract

This work proposes a novel learning-based method to address multimodal diffeomorphic registration. Traditional algorithms that address dense registration use numerical optimization solvers on intensity-based similarity metrics, so they work best in the monomodal setting. This work tackles this task by modality-agnostic descriptors which encode structural self-similarity and a Neural Ordinary Differential Equation (Neural ODE) encoding the dynamics of the estimated registration, achieving state-of-the-art smoothness and registration accuracy.

Introduction

Diffeomorphic registration is a task in medical image analysis whose goal is finding a dense non-rigid transformation that spatially aligns two given images, referred to as moving and fixed, respectively. Causes like disease or growth can cause deformations between scans, and a registration algorithm can compensate for these deformations to analyze anatomical variability. The target transformations to be estimated are thus assumed to be smooth with smooth inverse. Medical imaging applications would require these spatial transformations to be global diffeomorphisms, so registration algorithms can classify populations up to smooth and invertible transformations which accomplish: (1) Conservation of shape topology; (2) No introduction of sharp artifacts or collapsing voxels; (3) Invertibility with the same smoothness properties. Diffeomorphisms have traditionally been estimated by time-discretized variational optimization, geodesic shooting or stationary velocity fields.

Overall performance of monomodal registration algorithms by numerical optimization has been crucially limited by several factors: (1) Traditional methods assume the existence of local intensity self-similarity across fixed and moving images; (2) The underlying optimization incurs in high computational complexity, given by gradient computation over time-discrete velocity fields; (3) Existing biased anatomical similarity between the moving and the fixed inputs. Some methods have been proposed to integrate multimodal imaging sources, mainly by local descriptors on structural self-similarity or information theoretic similarity measures. Universal multimodal registration incurs an additional complexity gap given by a domain shift, in addition to the geometry shift present in the monomodal counterpart.

This work addresses pairwise multimodal registration driven by a structural similarity metric and leverages the deformation estimation by a Neural ODE method [1]. The inherent difficulty of registering varying-modality scans has conditioned state-of-the-art methods to perform domain transfer or neural optimal transport in the moving image domain, but they do not model accurate similarity metrics robust to structure and style divergence.

Method

Given a pair of fixed and moving images, deformable registration addresses the estimation of a diffeomorphism to warp the moving image into the fixed image by a variational optimization problem which depends on a similarity metric and achieves regularity and smoothness by weakly constraining the result with several regularizers (TV/L2 or H1 the most usual).

Pairwise registration by Neural ODEs

We encode the velocity field of the estimated transformations in a neural network, under the setting of a Neural ODE. This network outputs the velocity field of the desired transformations, and it is integrated forward or backward in time by an ODE solver. This ODE solver has been chosen to be Runge-Kutta or Euler in our experiments. We train for a maximum number of 300 iterations this network with functional regularization on the derivatives of the transformation's fields.

.

Structural Similarity

We encode voxel appearance in a descriptor based on the region self-similarity, to compare similar regions in terms of similar intensities, gradient and texture. The computation of this descriptor is performed by the MIND descriptor [2], so it permits to estimate appearance divergence by computing the sum of square differences (SSD) of the respective descriptors. For each voxel, we compute patch-wise descriptors according to a spatial scope of a 6-neighborhood. The result is scaled on the local variance of each patch and composed with an exponential kernel. This operation introduces nonlinearity and permits achieving higher signal response by self-similar regions of interest and robustness to noise.

Results

The non-rigid registration methods are evaluated according to Learn2Reg challenge measuring accuracy, orientation conservation and global smoothness. These three criteria are evaluated according to the mean Dice similarity coefficient (DSC), the sum of negative Jacobian determinants and the global ratio of negative Jacobians of the estimated transformations. The Dice score is computed on 32 segmentation labels.

Conclusions

We have introduced instance-specific methods on diffeomorphic registration which perform better quantitatively and qualitatively than other domain transfer strategies. These fail to capture richness on intra-subject variability across intensity dependence.

We evaluate this method with different functional regularizers and other structural metrics, as well as three domain transfer approaches. Our results are compared with two baselines: NiftyReg and ANTs. They perform registration by free-form deformations by dictionary learning on B-splines or by numerical optimization via information-theoretic measures (mutual information) respectively.

Acknowledgements

This work has received funding from Ministerio de Ciencia, Innovación y Universidades PID2022-138703OB-I00 and PROY-B50-24 granted by Gobierno de Aragón. Salvador Rodríguez-Sanz is sypported by a Gobierno de Aragón predoctoral grant (2024-2028).

References

[1] Yifan Wu, Tom Z Jiahao, Jiancong Wang, Paul A Yushkevich, M Ani Hsieh, and James C Gee. Nodeo: A neural ordinary differential equation based optimizationframework for deformable image registration. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 20804–20813, 2022

[2] Mattias P. Heinrich, Mark Jenkinson, Manav Bhushan, Tahreema Matin, Fergus V. Gleeson, Sir Michael Brady, and Julia A. Schnabel. Mind: Modality independent neighbourhood descriptor for multi-modal deformable registration. Medical Image Analysis, 16(7):1423–1435, 2012. ISSN 1361-8415.

T1-T2 OASIS-3 Test Set			
Baseline	DSC (%) $\uparrow$	Sum $\lvert J_\varphi \rvert \leq 0 \downarrow$	% of $\lvert J_\varphi \rvert \leq 0 \downarrow$
ANTs-SyN	68.75 ± 2.55	0	0
NiftyReg [44]	70.93 ± 1.50	0	0
NODE + Similarity $\mathcal{S}(\cdot,\cdot)$	DSC (%) $\uparrow$	Sum $\lvert J_\varphi \rvert \leq 0 \downarrow$	% of $\lvert J_\varphi \rvert \leq 0 \downarrow$
NODE + MIND [21]	$\mathbf{75.64 \pm 1.70}$	0	0
NODE-LDDMM + MIND [23]	$\mathbf{75.10 \pm 2.11}$	0	0
NODE + Local MI [46]	70.49 ± 3.80	0	0
NODE + Global MI [46]	44.90 ± 14.40	$2.94 \cdot 10^{-3}$	$1.01 \cdot 10^{-6}$
NODE + NGF [19]	72.40 ± 4.12	0	0

Figure 1. Quantitative results of the registration performance by the Neural ODE with modality-agnostic methods.

.