Extended Abstract Track

Rotation Equivariant Convolutions in Deformable Registration of Brain MRI

Editors: List of editors' names

Abstract

Image registration aligns anatomical structures between pairs of medical images. Deep learning-based registration methods have achieved state-of-the-art performance using convolutional neural networks (CNNs). However, while CNNs are translation equivariant, they lack rotation equivariance. This limitation prevents networks from fully exploiting the inherent rotational symmetries present in anatomical structures, particularly in brain MRI where these symmetries are prominent features. In this work, we investigate the impact of replacing standard convolutions with rotation equivariant convolutions in deformable brain MRI registration networks. We evaluate our approach on two baseline architectures (VoxelMorph and Dual-PRNet++) across multiple brain MRI datasets and compare against a non-symmetric control dataset. Our experiments demonstrate that rotation equivariant encoders improve registration accuracy on symmetric brain data while showing decreased performance on non-symmetric anatomical structures, confirming that the inductive bias of rotational symmetry is beneficial when anatomically justified.

Keywords: Image registration; CNN; Steerable kernel; rotation equivariance.

1. Introduction

Deformable image registration is the process of finding dense non-linear correspondences between pairs of fixed and moving images, serving as a cornerstone technique in medical image analysis. Traditional registration methods solve optimization problems iteratively for each image pair, making them computationally expensive. This led to the adoption of learning-based approaches that leverage CNNs to predict deformation fields directly from input pairs. Methods such as VoxelMorph Balakrishnan et al. (2019) and its variants have demonstrated that end-to-end learning can achieve comparable or superior performance to traditional optimization-based approaches while being orders of magnitude faster at inference time. Despite their success, current CNN-based registration methods do not fully exploit the geometric symmetries inherent in medical images. CNNs are translation equivariant by design; applying a spatial translation to the input produces a corresponding translation in the output feature maps. This property serves as a powerful inductive bias that promotes parameter sharing, reduces overfitting, and improves generalization. However, CNNs lack rotation equivariance. This limitation is particularly problematic for medical image registration, where anatomical structures often exhibit clear rotational symmetries. Brain MRI exemplifies this challenge, as neuroanatomical structures display prominent symmetry specially in the cerebral cortex.

Recent advances in geometric deep learning have introduced rotation-equivariant convolutions through steerable kernels Weiler and Cesa (2019); Cesa et al. (2022b), which constrain convolutional filters to transform predictably under rotations. These methods have shown promise in various computer vision tasks but remain underexplored in medical image registration.

In this work, we investigate the integration of rotation-equivariant convolutions into deformable brain MRI registration by replacing standard convolutional encoders in two representative architectures, VoxelMorph and Dual-PRNet++, with SE(3)-equivariant encoders that respect both rotational and translational symmetries. Our contributions are: (1) We demonstrate that rotation equivariance improves registration accuracy on symmetric brain MRI data while maintaining computational efficiency, (2) We validate our approach through controlled experiments on non-symmetric data (AbdomenCT) and show that benefits stem from exploiting relevant geometric symmetries rather than increased model capacity.

2. Related Work

2.1. Unsupervised Learning-based Registration

Deep learning has achieved state-of-the-art performance in medical image registration. Since obtaining ground-truth deformation fields is both time-consuming and often ambiguous, unsupervised registration methods have gained prominence, which optimize a loss function combining similarity and regularization terms.

VoxelMorph Balakrishnan et al. (2019) pioneered end-to-end learning with a U-Net-style architecture that directly predicts deformation fields, achieving competitive performance while being significantly faster than iterative methods. Multi-resolution approaches like LapIRN Mok and Chung (2020) employ cascaded frameworks that progressively refine deformations across resolution levels for better handling of large deformations. Recent works have explored diverse architectures including transformer-based methods (NIC-Trans Meng et al. (2023)), correspondence-based approaches like Dual-PRNet++ Kang et al. (2022) with dual-stream encoders and 3D correlation volumes. Additionally, RDP Wang et al. (2024) introduces recursive deformation pyramids for improved multi-scale registration.

Despite their success, existing methods rely on standard convolutions that lack rotation equivariance, potentially limiting performance when images exhibit orientational variations, particularly in anatomically symmetric regions like the brain.

2.2. Rotation Equivariant Networks and Steerable Kernels

Early Group-CNNs Cohen and Welling (2016a) achieved equivariance to finite transformation groups by creating multiple transformed filter versions, but suffered from memory overhead scaling with group size. Steerable CNNs Cohen and Welling (2016b) addressed this by using group representation theory to enable equivariance to continuous groups, constructing kernels with spherical harmonics and learnable radial functions for efficient SE(3) equivariance. Medical imaging applications have demonstrated benefits across various tasks. Winkels and Cohen (2018) applied group convolutions to chest X-ray analysis, while Diaz et al. (2024) showed improvements in 3D brain tumor and multi-organ segmentation. Moyer et al. (2021) and Billot et al. (2024) introduced SE(3) equivariance for rigid motion tracking in brain MRI.

Unlike previous works focusing on classification, segmentation, or rigid registration, we investigate rotation equivariant convolutions in deformable registration networks, where geometric inductive biases must balance against learning complex non-linear deformation fields.

3. Method

3.1. Problem Formulation

Given a moving and fixed image $I_m, I_f \in \mathbb{R}^{H \times W \times D}$, medical image registration aims to find a dense deformation field $\phi : \Omega \to \Omega$ to align the anatomical structures of the warped moving image $I_m \circ \phi$ with the fixed image I_f, where $\circ$ denotes spatial transformation. The registration process minimizes a loss function of the form: $\mathcal{L}(\phi) = \mathcal{L}_{\text{sim}}(I_f, I_m \circ \phi) + \lambda \mathcal{L}_{\text{reg}}(\phi)$ where L_{sim} measures the similarity between the aligned images, L_{reg} enforces spatial smoothness of the deformation field and λ is a regularization weight.

3.2. Baseline Architectures

We evaluate rotation equivariance on two representative architectures by replacing their standard encoders with equivariant encoders while keeping decoders unchanged to isolate the effect of equivariant feature extraction: **(1)VoxelMorph:** A single-stage U-Net that takes concatenated images as input and directly predicts deformation fields Balakrishnan et al. (2019) **(2)Dual-PRNet++:** A multi-stage architecture with dual-stream encoders and pyramid registration using 3D correlation volumes Kang et al. (2022).

These architectures differ in processing strategy (single vs. multi-stage), input handling (joint vs. separate moving and fixed images), and complexity (simple U-Net vs. pyramid-based modules), ensuring our findings generalize across different registration frameworks.

3.3. Steerable Kernels

Standard convolutions operate on scalar features, limiting their ability to capture geometric relationships fundamental to registration. We employ steerable kernels that operate on higher-order geometric features with well-defined transformation properties under rotations.

Any finite-dimensional representation of SO(3) can be decomposed into irreducible representations (irreps) indexed by $l = 0, 1, 2, ...$, with dimension $d = 2l + 1$. We use scalars ($l = 0$) and vectors ($l = 1$) in our experiments.

A key contribution of Weiler et al. (2018) was the derivation of a constraint governing kernels between irreps of orders l and l', and the corresponding construction of a complete basis for these steerable kernels.

3.4. Equivariant Encoder Design

For each baseline, we design SE(3)-equivariant encoders that replace standard encoders while preserving decoders, isolating the impact of equivariant feature extraction. Each block uses SE(3)-steerable convolution with gated activation Weiler et al. (2018), matching architectural parameters to baseline models. All layers use 50% irrep(0) + 50% irrep(1), capturing rotation-invariant and directional features efficiently. the channel dimension here may increase, but the total number of parameters are still comparable.

4. Experiments

We conducted our experiments on two main 3D brain MRI datasets: **(1)OASIS** Marcus et al. (2007) with 35 anatomical labels. We used 200 brain MRI scans for training, 20 for validation, and the provided test set for evaluation. **(2)LPBA40** Shattuck et al. (2008) with 54 anatomical labels. We used 25 samples for training, 5 for validation, and 10 for testing. We also used the **AbdomenCT** Xu et al. (2016) dataset as a control experiment to compare the effect of equivariant encoders on data that does not exhibit the rotational symmetries. We have implemented the models in PyTorch, and for steerable kernels, we have utilized the escnn library Cesa et al. (2022a).

5. Results

	OASIS		**LPBA**		**Abdomen CT**	
	DSC ↑	**ASSD** ↓	**DSC** ↑	**ASSD** ↓	**DSC** ↑	**ASSD** ↓
VM	0.794 ± 0.025	0.709 ± 0.108	0.658 ± 0.028	2.008 ± 0.216	0.375 ± 0.051	4.598 ± 0.958
Equi VM	0.794 ± 0.025	0.710 ± 0.111	0.660 ± 0.030	1.997 ± 0.226	0.371 ± 0.051	4.647 ± 0.961
Dual PR ++	0.817 ± 0.020	0.630 ± 0.089	0.667 ± 0.030	1.964 ± 0.225	0.466 ± 0.056	3.883 ± 0.929
Equi Dual PR ++	0.816 ± 0.021	0.632 ± 0.088	0.688 ± 0.023	1.841 ± 0.180	0.451 ± 0.059	4.107 ± 0.960

Table 1: Comparison of DSC (Dice Similarity Coefficient) and ASSD (Average Symmetric Surface Distance) across OASIS, LPBA, and Abdomen CT datasets.

We evaluated registration performance on two symmetric datasets (brain MRI) and one non-symmetric dataset (AbdomenCT) using both baseline models and their corresponding equivariant variants with irreps(0) and (1).

Results in Table 1 show that equivariant encoders either improve (LPBA) or match (OASIS) performance on brain datasets, while decreasing performance on AbdomenCT data. These findings align with our hypothesis that equivariant encoders improve feature extraction through rotational symmetry inductive bias when anatomically justified, but degrade performance when this assumption is violated, as observed in the AbdomenCT dataset.

6. Discussion

In this work, we investigated the effectiveness of employing rotation equivariant convolutions in deformable brain MRI registration. Compared to non-symmetric data, we observed improved or maintained performance in equivariant models for symmetric brain data, confirming that geometric inductive biases provide benefits when anatomically justified. However, our experiments remain limited in scope and require expansion to additional registration architectures and diverse anatomical datasets. Furthermore, exploration of higher-order irreducible representations beyond scalars and vectors represents a promising direction for future work to capture more complex geometric relationships in anatomical structures.

References

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*, 38(8):1788–1800, 2019.

Benjamin Billot, Neel Dey, Daniel Moyer, Malte Hoffmann, Esra Abaci Turk, Borjan Gagoski, P Ellen Grant, and Polina Golland. Se (3)-equivariant and noise-invariant 3d rigid motion tracking in brain mri. *IEEE transactions on medical imaging*, 43(11): 4029–4040, 2024.

Gabriele Cesa, Leon Lang, and Maurice Weiler. A program to build E(N)-equivariant steerable CNNs. In *International Conference on Learning Representations*, 2022a. URL https://openreview.net/forum?id=WE4qe9xlnQw.

Gabriele Cesa, Leon Lang, and Maurice Weiler. A program to build e (n)-equivariant steerable cnns. In *International conference on learning representations*, 2022b.

Taco Cohen and Max Welling. Group equivariant convolutional networks. In *International conference on machine learning*, pages 2990–2999. PMLR, 2016a.

Taco S Cohen and Max Welling. Steerable cnns. *arXiv preprint arXiv:1612.08498*, 2016b.

Ivan Diaz, Mario Geiger, and Richard Iain McKinley. Leveraging so (3)-steerable convolutions for pose-robust semantic segmentation in 3d medical data. *The journal of machine learning for biomedical imaging*, 2(May 2024):834, 2024.

Miao Kang, Xiaojun Hu, Weilin Huang, Matthew R Scott, and Mauricio Reyes. Dual-stream pyramid registration network. *Medical image analysis*, 78:102379, 2022.

Daniel S Marcus, Tracy H Wang, Jamie Parker, John G Csernansky, John C Morris, and Randy L Buckner. Open access series of imaging studies (oasis): cross-sectional mri data in young, middle aged, nondemented, and demented older adults. *Journal of cognitive neuroscience*, 19(9):1498–1507, 2007.

Mingyuan Meng, Lei Bi, Michael Fulham, Dagan Feng, and Jinman Kim. Non-iterative coarse-to-fine transformer networks for joint affine and deformable image registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 750–760. Springer, 2023.

Tony CW Mok and Albert CS Chung. Large deformation diffeomorphic image registration with laplacian pyramid networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 211–221. Springer, 2020.

Daniel Moyer, Esra Abaci Turk, P Ellen Grant, William M Wells, and Polina Golland. Equivariant filters for efficient tracking in 3d imaging. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 193–202. Springer, 2021.

David W Shattuck, Mubeena Mirza, Vitria Adisetiyo, Cornelius Hojatkashani, Georges Salamon, Katherine L Narr, Russell A Poldrack, Robert M Bilder, and Arthur W Toga. Construction of a 3d probabilistic atlas of human cortical structures. *Neuroimage*, 39(3): 1064–1080, 2008.

Haiqiao Wang, Dong Ni, and Yi Wang. Recursive deformable pyramid network for unsupervised medical image registration. *IEEE Transactions on Medical Imaging*, 43(6): 2229–2240, 2024.

Maurice Weiler and Gabriele Cesa. General e (2)-equivariant steerable cnns. *Advances in neural information processing systems*, 32, 2019.

Maurice Weiler, Mario Geiger, Max Welling, Wouter Boomsma, and Taco S Cohen. 3d steerable cnns: Learning rotationally equivariant features in volumetric data. *Advances in Neural information processing systems*, 31, 2018.

Marysia Winkels and Taco S Cohen. 3d g-cnns for pulmonary nodule detection. *arXiv preprint arXiv:1804.04656*, 2018.

Zhoubing Xu, Christopher P Lee, Mattias P Heinrich, Marc Modat, Daniel Rueckert, Sebastien Ourselin, Richard G Abramson, and Bennett A Landman. Evaluation of six registration methods for the human abdomen on clinically acquired ct. *IEEE Transactions on Biomedical Engineering*, 63(8):1563–1572, 2016.