

Unified Brain Surface and Volume Registration

S. MAZDAK ABULNAGA, MIT CSAIL and MGH, Harvard Medical School, USA

ANDREW HOOPES, MIT CSAIL and MGH, Harvard Medical School, USA

MALTE HOFFMANN, MGH, Harvard Medical School

ROBIN MAGNET, Université Paris Cité, INRIA, France

MAKS OVSIJANIKOV, LIX, CNRS, École Polytechnique, France

LILLA ZÖLLEI, MGH, Harvard Medical School, USA

JOHN GUTTAG, MIT CSAIL, USA

BRUCE FISCHL*, MGH, Harvard Medical School, USA

ADRIAN V. DALCA*, MIT CSAIL and MGH, Harvard Medical School, USA

Accurate registration of brain MRI scans is fundamental for cross-subject analysis in neuroscientific studies. This involves aligning both the cortical surface of the brain and the interior volume. Traditional methods treat volumetric and surface-based registration separately, which often leads to inconsistencies that limit downstream analyses. We propose a deep learning framework, NeurAlign, that registers 3D brain MRI images by jointly aligning both cortical and subcortical regions through a unified volume-and-surface-based representation. Our approach leverages an intermediate spherical coordinate space to bridge anatomical surface topology with volumetric anatomy, enabling consistent and anatomically accurate alignment. By integrating spherical registration into the learning, our method ensures geometric coherence between volume and surface domains. In a series of experiments on both in-domain and out-of-domain datasets, our method consistently outperforms both classical and machine learning-based registration methods—improving the Dice score by up to 7 points while maintaining regular deformation fields. Additionally, it is orders of magnitude faster than the standard method for this task, and is simpler to use because it requires no additional inputs beyond an MRI scan. With its superior accuracy, fast inference, and ease of use, NeurAlign sets a new standard for joint cortical and subcortical registration.

ACM Reference Format:

S. Mazdak Abulnaga, Andrew Hoopes, Malte Hoffmann, Robin Magnet, Maks Ovsjanikov, Lilla Zöllei, John Guttag, Bruce Fischl, and Adrian V. Dalca. 2026. Unified Brain Surface and Volume Registration. 1, 1 (June 2026), 12 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

*Co-senior authors.

Authors' Contact Information: S. Mazdak Abulnaga, MIT CSAIL and MGH, Harvard Medical School, USA, abulnaga@csail.mit.edu; Andrew Hoopes, MIT CSAIL and MGH, Harvard Medical School, USA; Malte Hoffmann, MGH, Harvard Medical School; Robin Magnet, Université Paris Cité, INRIA, France; Maks Ovsjanikov, LIX, CNRS, École Polytechnique, France; Lilla Zöllei, MGH, Harvard Medical School, USA; John Guttag, MIT CSAIL, USA; Bruce Fischl, MGH, Harvard Medical School, USA; Adrian V. Dalca, MIT CSAIL and MGH, Harvard Medical School, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM XXXX-XXXX/2026/6-ART

<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 Introduction

Accurate alignment of both cortical *and* subcortical regions is essential for comprehensive and consistent whole-brain analysis in neuroimaging studies. For example, functional MRI (fMRI) measures brain activity by detecting blood oxygenation changes, with a key goal of localizing functional regions in the cerebral cortex. Mapping activation patterns to brain structure reveals functional organization and is essential for quantifying how neurodegenerative diseases affect things such as cognition [23, 24, 56].

Traditional deformable volumetric registration methods compute a three-dimensional (3D) displacement field to align brain images by maximizing intensity similarity [5, 7, 49]. While effective for aligning subcortical structures and global anatomy, volumetric deformable registration often fails in the cortex. The cortex is a thin, highly curved surface with significant inter-subject variability in folding patterns that is difficult to align in Euclidean space [22]. Accurate alignment is critical, since the cortex encodes many of the brain's high-level functions such as memory and language, and serves as a primary target for structural and functional neuroimaging studies [21]. Surface-based methods address this challenge by projecting the cortex onto a sphere and aligning the cortex using pre-computed anatomical features such as sulcal depth and/or curvature. This facilitates registration by providing a common geometric space to preserve the topology of the cortex [22, 61].

Using fundamentally different representations for surface and volume registration forces neuroscience researchers to tackle two disjoint problems and combine solutions *ad hoc* [58]. This typically involves solving an elastic partial differential equation to propagate the surface registration to the interior brain. However, interpolating the surface registration to the interior is not guaranteed to be consistent with the original volumetric registration. This sequential approach prevents computing a coherent registration that simultaneously satisfies both cortical and subcortical alignment objectives. This introduces errors, undermining whole-brain analyses by potentially misaligning anatomically and functionally connected areas [3, 33, 44].

In this work, we propose NeurAlign, a machine learning framework for unified registration of cortical and subcortical structures. NeurAlign, simultaneously trains a volumetric registration network and a spherical registration network using a shared framework that

combines both domains. Unlike existing approaches, our model explicitly couples the two through a loss function that encourages the volumetric deformation field on the cortical ribbon to match the spherical registration field. This promotes topologically correct and geometrically consistent alignment across cortical and subcortical regions. Crucially, because our method performs registration in a single forward pass, it enables efficient and scalable whole-brain registration for large population studies. Further, NeurAlign does not require cortical meshes nor segmentations at inference, avoiding the computationally intensive preprocessing required by other methods. Our method, NeurAlign, achieves significantly improved cortical and subcortical alignment compared to the standard joint registration framework, CVS [44], while reducing computation time by several orders of magnitude. In addition, NeurAlign substantially outperforms state-of-the-art volumetric machine learning-based registration approaches. To summarize our contributions:

- We propose a unified machine learning model that performs consistent registration of both cortical and subcortical structures using coupled spherical and volumetric networks.
- We derive a cortical consistency loss that explicitly encourages agreement between the volumetric deformation field in the cortex and the spherical registration field, promoting anatomically coherent alignment across the entire brain.
- Our proposed method leverages cortical meshes during *training*, but does not require them during inference, enabling accessible, rapid, and accurate registration.
- We validate our method on four clinical neuroimaging datasets, showing substantially improved and rapid cortical and subcortical alignment compared to standard baselines.

Our code and model weights are available at <https://github.com/mabulnaga/neuralign>.

2 Related Work

Volumetric Registration. Deformable registration estimates a dense spatial mapping between a pair of images. Classical methods [4, 40, 47, 49] solve an optimization problem for each pair of images. This balances an image-similarity objective, often mean-squared error (MSE) or normalized cross correlation (NCC) [5], between the warped image and the fixed image, with a regularization objective that encourages a smooth solution. Methods differ in their optimization strategy, choice of the objective function, and in how they parameterize the deformation field.

Learning-based methods fit the parameters of a neural network to a training dataset. The network learns a nonlinear function that maps an input image pair to an output transform, generally enabling much faster inference than classical methods. Supervised approaches [46, 53, 60] reproduce known simulated deformations or fields estimated by classical methods, whereas unsupervised approaches [2, 7, 17, 25, 26, 41, 45] optimize an image-similarity loss term and a regularization loss term at training, analogous to classical registration.

Both optimization-based and machine learning-based volumetric approaches are highly effective at aligning subcortical structures. However, they struggle with cortical registration, since the optimization landscape in Euclidean space often leads to suboptimal local

minima. This can lead to implausible solutions that, for example, move voxels from one gyrus to the next instead of following the cortical fold. Further, relying only on image-based features cannot align the functional regions of the cortex, so geometric features must be employed from geometric representations [22, 51].

Spherical Registration. Cortical alignment in the spherical domain is the standard approach in neuroscientific studies [21, 24, 58]. Mapping the cortex to the sphere simplifies the optimization, since the registration is guaranteed to preserve topology. The registration is driven by geometric descriptors that measure the cortical shape and folding patterns [14, 22, 61, 62]. Similar to volumetric registration, spherical registration is posed as an optimization problem that balances a geometric similarity objective with a regularization objective. Recent methods use deep learning to improve inference time [13, 37, 62]. The spherical mesh is parameterized to the 2D plane using a stereographic projection. The registration is performed in 2D space using convolutional neural networks [62]. Icosahedral CNNs [63] learn directly on the sphere but require a rigid tessellation that is not produced natively by standard neuroimaging pipelines. These methods have achieved comparable accuracy to classical techniques with a significant improvement in speed. We build on the planar parameterization methods to construct a spherical registration network and jointly model it with a volumetric network.

Surface Mesh Registration. Surface registration aligns two meshes or point clouds in 3D Euclidean space. Classical approaches typically frame the task as a deformation problem, aiming to align two shapes by estimating a rigid or non-rigid transformation, that minimizes a distance metric [1, 8, 50, 54, 59]. An alternative paradigm treats surface registration as a *shape correspondence* problem, focusing on establishing point-to-point correspondences based on intrinsic geometry [18]. Functional maps [43] and extensions [12, 20, 52] have emerged as a powerful representation for such correspondences, enabling compact linear descriptions of dense mappings between shapes. However these methods, developed for graphics applications target moderate complexity meshes and scale poorly to high-curvature, high-resolution cortical surfaces, despite some efforts for memory-efficient variants [38]. Furthermore, because of the high curvature and topological consistency constraints of cortical surfaces, Euclidean 3D embeddings become unstable and less informative. Finally, these methods are unsuitable for registering volumetric structures as they are only defined for 2-dimensional surfaces. These reasons further motivate the use of spherical registration to align cortical surfaces.

Combined Volume and Surface Registration. Existing neuroscientific studies treat volumetric and cortical registration as two separate problems. The Combined Volume and Surface framework (CVS) [44, 64] is the standard method for jointly aligning cortical and subcortical anatomy in adult brains. CVS first aligns the cortical surface then propagates the resulting deformation into the volume using an elastic model, followed by intensity-based refinement. While effective, this sequential formulation decouples the surface and volumetric objectives, can introduce inconsistencies near the cortical-subcortical boundary, and requires surface meshes

and long runtimes. Related work has also used cortical surface registration to supervise volumetric registration [3], though focused in the infant-brain domain. A strength of these variational approaches is that they offer a principled way to propagate surface constraints into the volume using well-understood physical models. However, they remain impractical for large-scale datasets, as they require several hours of computation per pair and depend on extracting cortical surfaces and segmentation labelmaps.

To address these limitations, we introduce NeurAlign, which incorporates the established idea of surface-guided alignment into a single, unified learning-based diffeomorphic framework. NeurAlign jointly aligns cortical and subcortical structures through a single forward pass, requires only 3D MRI at inference time (without surface meshes), and provides the speed and simplicity needed for high-throughput neuroimaging studies.

3 Methods

We develop NeurAlign, a method for jointly aligning cortical and subcortical structures in 3D brain MRI. We jointly optimize a volumetric and a spherical deformation field and propose a soft constraint to encourage consistency between the two registration fields. We first propose the continuous formulation of the problem before describing the discrete case.

3.1 Continuous Model

Let $M_1, M_2 \subset \mathbb{R}^3$ be bounded brain volumes, with associated smooth cortical surfaces $\partial M_1, \partial M_2$. We seek a bijective map $\varphi : M_1 \rightarrow M_2$ that aligns both the cortical and subcortical structures. Each volume is equipped with an intensity sampling function $I_i : \mathbb{R}^3 \rightarrow \mathbb{R}$. The surfaces are also equipped with d geometric descriptor functions $g_i : \partial M_i \rightarrow \mathbb{R}^d$, such as sulcal depth or mean curvature.

Finding φ amounts to finding a solution to the following optimization problem,

$$\arg \min_{\varphi \in \mathcal{P}} \int_{M_1} f_I(I_1(\mathbf{x}), I_2(\varphi(\mathbf{x}))) dV(\mathbf{x}) + \int_{\partial M_1} f_S(g_1(\mathbf{x}), g_2(\varphi(\mathbf{x}))) dS(\mathbf{x}) + R[\varphi], \quad (1)$$

with $\mathcal{P}$ denoting the feasible set of maps, enforcing for instance $\varphi(\partial M_1) \subset \partial M_2$.

Here, f_I measures image similarity, f_S compares surface descriptors, dV and dS denote the volume and surface measures, and $R[\cdot]$ denotes a set of regularization functionals.

Solving Eq. (1) is difficult: the joint matching of intensities and geometric features has a nonconvex objective with degenerate solutions. Enforcing $\varphi(\partial M_1) \subseteq \partial M_2$ is especially challenging as the cortex is thin, highly folded, and has variable structure across subjects. Fine-grained boundary correspondence is poorly captured in a Euclidean volumetric grid, and optimizers often only align subsets of folds while blurring or misaligning others [51]. Finally, finding a unified representation that captures both the volumetric image and cortical surface is difficult.

We use a joint representation, where cortical alignment is performed on the sphere, using topology-preserving maps and volumetric registration is performed in 3D. By enforcing consistency

between the two, this learning-based formulation efficiently registers both the cortical and subcortical structures.

3.2 Spherical Alignment

We revise the optimization problem in Eq. (1) by introducing a spherical intermediate domain to help enforce the constraint set $\mathcal{P}$. Let $\tau_i : \partial M_i \rightarrow \mathbb{S}^2$ denote the (fixed) invertible spherical maps obtained via cortical inflation [22, 30] for each surface. The cortical alignment problem then computes a map $\psi : \mathbb{S}^2 \rightarrow \mathbb{S}^2$ that displaces spherical coordinates (θ, ϕ) , to align geometric descriptors. By optimizing on the sphere, we guarantee satisfaction of the constraint $\mathcal{P}$. The revised optimization problem is then:

$$\arg \min_{\varphi, \psi \in \mathcal{Q}} \int_{M_1} f_I(I_1(\mathbf{x}), I_2(\varphi(\mathbf{x}))) dV(\mathbf{x}) + \int_{\partial M_1} f_S(g_1(\tau(\mathbf{x})), g_2(\psi(\tau(\mathbf{x})))) dS(\tau(\mathbf{x})) + R[\varphi] + R[\psi], \quad (2)$$

We decouple this optimization into two separate optimization problems: volumetric registration of subcortical regions using φ , and cortical registration on the sphere using ψ . We introduce a constraint set $\mathcal{Q}$ to encourage a consistent mapping below.

Cortical Alignment Consistency Constraint. Optimizing Eq. (2) independently can estimate maps φ and ψ that map the cortex to different locations in the volume. We therefore require consistency in Euclidean space by composing maps to the sphere and back: $\mathcal{Q} = \{\varphi(\mathbf{x}) = \tau_2^{-1} \circ \psi \circ \tau_1(\mathbf{x}) \text{ for } \mathbf{x} \in \partial M_1\}$.

To keep the problem tractable, we soften this constraint using a coupling energy that encourages consistent alignment:

$$E_{cons}(\varphi, \psi) = \int_{\partial M_1} f(\varphi(\mathbf{x}), \tau_2^{-1}(\psi(\tau_1(\mathbf{x})))) dS(\mathbf{x}), \quad (3)$$

where f measures dissimilarity in the mapping by the squared error. We replace the hard constraint $\mathcal{Q}$ with the coupling energy (Eq. (3)) to achieve a consistent mapping that accurately aligns both cortical and subcortical structures. This constraint only acts on a set of Lebesgue measure zero in $\mathbb{R}^3$ (the cortical surface ∂M_1) and is therefore weak. In practice, however, our discrete implementation avoids this issue: the consistency loss samples the volumetric deformation at mesh vertex locations via trilinear interpolation, distributing gradients to neighboring voxels. We also impose smoothness priors on the map, allowing the surface-derived updates to naturally propagate throughout the volume.

The remainder of this section focuses on the discretization and use of neural networks to model the registration problem.

3.3 Discrete Model

We discretize the optimization problem in Eq. (2), and use neural networks to parameterize the volumetric and spherical maps. Specifically, we formulate an unsupervised learning problem where we jointly learn 3D and 2D U-Net convolutional neural networks (CNN) [48]. Given an image and a surface, the networks output the 3D and 2D deformation fields φ, ψ , respectively. We propose a consistency loss to obtain an accurate cortical surface registration in the volume. Figure 1 presents an overview of our framework.

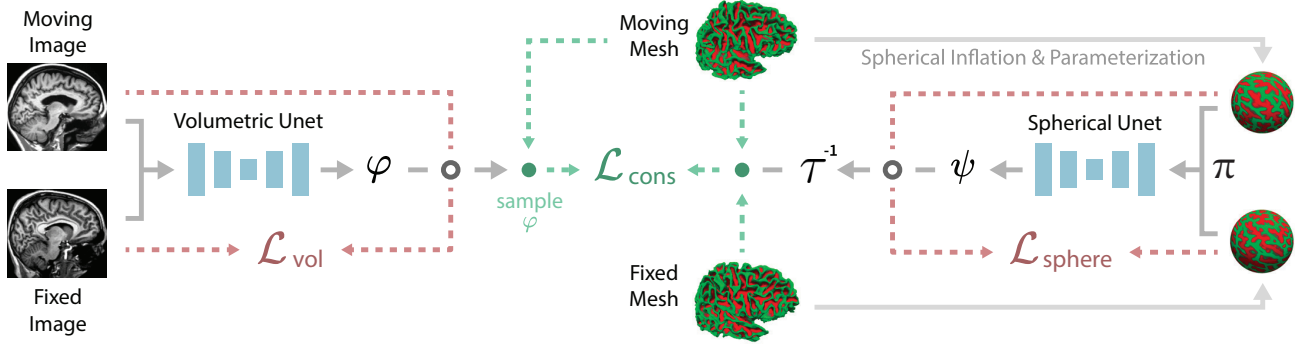


Fig. 1. Unified cortical and volumetric registration framework. We employ a volumetric CNN to perform 3D image alignment, and a 2D spherical CNN to align the cortical surfaces in the spherical domain. In training, the models require the MRI images, the cortical surface meshes, and the inflated spheres. The resultant deformation fields of both models are used in a loss term ($\mathcal{L}_{cons}$) that encourages cortical registration to be consistent in the spherical and volumetric pathways. At inference time, the 3D model does not require meshes nor spheres to perform registration.

Notation. Let $I_1, I_2 : \mathbb{R}^3 \rightarrow \mathbb{R}$ denote two 3D brain MRI scans, where $I_i(\mathbf{x})$ gives the intensity at spatial location $\mathbf{x}$. Each image used in training the networks also has an associated triangular surface mesh $C_1, C_2 \subset \mathbb{R}^3$ representing the cortical surface extracted from the respective anatomy. Vertex j of mesh i , v_j^i , is a 3D coordinate in the image space of I_i . Each cortical surface mesh has an associated mesh in the spherical domain denoted S_1, S_2 , where $S_1 = \tau_1(C_1)$, with the same triangle connectivity. The vertices v are mapped to the sphere by a standard spherical inflation procedure and have an associated $\mathbb{S}^2$ representation [22].

Volumetric Registration Network. We employ a neural network $F_v(I_1, I_2; \omega_v) = \varphi$ to predict φ as a function of the input images, where F_v is parameterized by ω_v . A core 3D convolutional network first outputs a stationary velocity field (SVF), which we then integrate to obtain a diffeomorphic displacement field φ [4, 15]. In the deformation field $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, each voxel in $\mathbb{R}^3$ is assigned a displacement vector $\mathbf{u} : \varphi(\mathbf{x}) = \mathbf{x} + \mathbf{u}(\mathbf{x})$.

Spherical Registration Network. We construct a 2D CNN registration network for the sphere. We parameterize the spherical meshes to the 2D Euclidean plane using a stereographic projection $\pi : \mathbb{S}^2 \rightarrow \mathbb{R}^2$. By mapping to the plane, we can use a standard 2D CNN to learn a 2D diffeomorphic displacement field in the space of spherical angles (θ, ϕ) .

To account for the nonuniform sampling of the spherical parameterization, where regions near the poles are stretched, we apply distortion correction in all surface-based losses by weighting samples by $\sin(\theta)$, where θ is the elevation angle [13, 37]. We also handle the discontinuities at the image boundaries using the padding strategy of [37], employing circular padding along the left–right axis and a 180° circular shift with reflection padding at the poles.

We use a neural network $F_s(\pi(S_1), \pi(S_2); \omega_s) = \psi$ to predict ψ , by first predicting a SVF that is integrated. The network outputs a displacement field $\mathbf{u}_s$ such that $\psi(\rho) = \rho + \mathbf{u}_s(\rho)$, where $\rho = (\theta, \phi)$ is a discrete location in pixel coordinates on the parameterized grid.

3.3.1 Unified Training and Cortical Alignment.

Cortical Consistency Loss. We maximize alignment of the cortex and subcortical structures between a pair of images while maintaining a smooth map.

We discretize the proposed consistency loss in Eq. (3), and use it to jointly train the volumetric and spherical registration networks. The consistency loss penalizes disagreement between the predicted 3D displacement on the cortex, and the 3D displacement computed from the predicted 2D warp:

$$\mathcal{L}_{cons}(\varphi, \psi, C_1) = \frac{1}{N_v} \sum_{v \in C_1} f(\varphi(v), \pi^{-1}(\psi(\pi(v)))), \quad (4)$$

where N_v is the number of vertices in C_1 , $\pi(v)$ is the projection of v to the plane, and $f(\cdot, \cdot)$ is a similarity metric; we use the mean squared error in our experiments. We implement $\pi^{-1}(\cdot)$ with trilinear interpolation over the mapped planar parameterizations of v . Maps φ and ψ are the outputs of the 3D and 2D CNN, respectively.

Image Similarity Loss. Both networks are trained by minimizing an unsupervised loss that balances image-similarity and field regularity. The similarity loss on the volume matches MRI intensities and the loss on the sphere matches geometric descriptors. We precompute geometric descriptors on the cortical meshes, $g : C \rightarrow \mathbb{R}^2$. As is standard in spherical registration, we use the sulcal depth and mean curvature [13, 22]. The similarity loss is $\mathcal{L}_{sim}(\varphi, \psi) = \mathcal{L}_{sim}(I_2, \varphi(I_1)) + \mathcal{L}_{sim}(g(\pi(S_2)), \psi(g(\pi(S_1))))$. Here, $w(\rho)$ is a weighting to correct for area distortion introduced by the stereographic projection. The similarity loss measures pairwise similarity between the fixed image and sphere I_2, S_2 and the mapped image and sphere, $\varphi(I_1), \psi(S_1)$. We use the local normalized cross correlation function as our similarity measure [6].

Regularization. We regularize the displacement fields to be smooth to encourage a well-behaved map. Our regularization loss is

$$\mathcal{L}_{reg}(\varphi, \psi) = \|\nabla \varphi\|^2 + \|\nabla \psi\|^2,$$

where $\nabla \varphi$ is the Jacobian of map φ .

Auxiliary structural loss. The use of anatomical label maps during training of learning-based registration often improves substructure alignment [7]. When segmentation maps are available for some

pairs of images, we use this information to supervise by structural alignment. Let $\text{seg}[I_1]$ indicate the segmentation labelmap of the K brain substructures for image I_1 . We add an additional loss to align these structures, $\mathcal{L}_{\text{struc}}(\varphi, \psi) = \mathcal{L}_{\text{struc}}(\text{seg}[I_2], \varphi(\text{seg}[I_1])) + \mathcal{L}_{\text{struc}}(\pi(\text{seg}[\mathcal{S}_2]), \pi(\psi(\text{seg}[\mathcal{S}_2])))$. In practice, we use the soft Dice loss function, which captures volume overlap of the warped subcortical structures on one image warped to the second.

Joint training loss. Combining all loss terms provides our joint optimization problem:

$$\mathcal{L}(\varphi, \psi) = \mathcal{L}_{\text{sim}}(\varphi, \psi) + \gamma \mathcal{L}_{\text{cons}}(\varphi, \psi) + \lambda \mathcal{L}_{\text{reg}}(\varphi, \psi) + \kappa \mathcal{L}_{\text{struc}}(\varphi, \psi), \quad (5)$$

where λ , γ , and κ are hyperparameters.

4 Experiments

We evaluate NeurAlign using 3D brain MRI scans from multiple datasets comprising healthy subjects and subjects with cognitive disorders. We compare our method against CVS, the state-of-the-art combined cortical surface-volume registration framework [44] and against learning-based methods that emphasize structural alignment. We test the ability of our method to generalize to new datasets unseen in training.

4.1 Experimental Setup

Data. We train our model using two public 3D brain MRI datasets: OASIS-1 [39] and ADNI [42]. Both datasets are obtained from large population studies focused on Alzheimer’s disease (AD). OASIS-1 [39] includes T1-weighted (T1w) scans of 416 subjects aged 18–96. We use a random subset of 100 ADNI subjects aged 56–90, half of whom were diagnosed with AD. For each ADNI subject, we use data from two sessions, two years apart.

For each dataset, we hold out 20% of the subjects for testing, stratified by gender, age, and disease state. The remaining subjects are split into 85% for training and 15% for validation using the same stratification. In training, we randomly sample pairs from both datasets. In our test set, we create 123 non-overlapping pairs for registration: 83 pairs aligning OASIS-1 and OASIS-1, 20 pairs aligning ADNI and ADNI, and 20 pairs aligning OASIS-1 and ADNI. **Held-out datasets.** We use two additional held-out datasets for testing model generalization. The IXI dataset contains MRI scans from healthy subjects acquired in 3 London hospitals [31]. We randomly select T1w scans from 115 subjects for evaluation. We also hold out 80 T1w scans from the Mindboggle-101 dataset [35], a collection of manually labeled brain MRI volumes and surfaces. We exclude all OASIS-1 subjects from the Mindboggle dataset.

All evaluations are performed on T1w MRI, which is the modality that enables reliable cortical surface extraction. T1w MRI has the strongest gray-white matter contrast and supports reliable delineation of white and pial surfaces [21, 22, 30].

Processing. We use FreeSurfer [21] to perform all image preprocessing. FreeSurfer is an open-source, widely used toolkit in neuroimaging that supports cortical surface analysis, longitudinal workflows, and population-scale MRI studies. We first perform affine alignment of each image to the Talairach template [55]. We crop all images to a roughly 10 mm margin around the brain. To prepare our training data, we generate anatomical segmentations, white matter (internal

and pial (external) cortical surface reconstructions, and spherically inflated surfaces with parameterized curvature maps. For evaluation, we use both volumetric segmentations of subcortical brain regions, computed on the volume, as well as parcellations of the cortical ribbon, computed on the surface mesh. Surface parcellations are rasterized as volumetric segmentations. We emphasize that our method does not require surface meshes, spheres, nor segmentation labelmaps at inference.

Implementation Details. We train using the Adam optimizer [34] with a learning rate of 10^{-4} . We set the field regularization hyperparameter to $\lambda = 1.0$, the segmentation loss weight to $\kappa = 10.0$, and the cortical consistency weight to $\gamma = 0.05$. These were chosen using a grid search over the validation set. We train for 300,000 iterations, using the best performing model on the validation set. All models are trained on a single A100 GPU. All CVS experiments are done on an Intel(R) Xeon(R) Gold 5218 CPU, as there is no GPU implementation.

Baselines. We select state-of-the-art baselines that emphasize structural alignment. We evaluate CVS [44, 64], the most widely used joint volume-cortical alignment registration method, which is integrated in the FreeSurfer software suite. CVS performs sequential spherical to volumetric alignment using a biomechanical model and image intensity registration.

We evaluate uGradICON [57], a modern foundation model for medical image registration that generalizes well to unseen data. UniGradICON employs test-time optimization to improve pairwise-registration accuracy. We also use the uGradICON-seg variant that performs loss function masking using the segmentation labelmaps to improve cortical structure alignment.

We test SynthMorph [27, 28], a state-of-the-art registration tool trained exclusively on images synthesized from label maps. Since this tool optimized the overlap of standard FreeSurfer labels in training, which include only two large cortical structures, we also retrain it with extended label sets, optimizing overlap of the “a2009s” set of 196 labels that include finer Destrieux parcels of the cortex [19] (SynthMorph-wm).

We evaluate VoxelMorph [16], an unsupervised diffeomorphic registration model. We use the same size UNet as our model and train on our training set. We also evaluate FiReANTs [32], a library for fast medical image registration using Riemannian optimization. We use the SyN transformation for symmetric diffeomorphic registration.

Evaluation. To evaluate registration accuracy, we compute Dice scores between anatomical structures in the warped and target images. We report overlap for 43 subcortical structures and for cortical regions obtained by parcellating the cortical ribbon into 34 regions per hemisphere. We assess field regularity and topology by computing the determinant of the Jacobian of the map, $\det J_\varphi(p) = \det(\nabla \varphi(p))$ at each voxel p . Locations where $\det J_\varphi(p) < 0$ represent folded regions. We report the percentage of folds (% folds). We measure the smoothness of the warp fields using the standard deviation of $\log |\det J|$ (SDlogJ) [36].

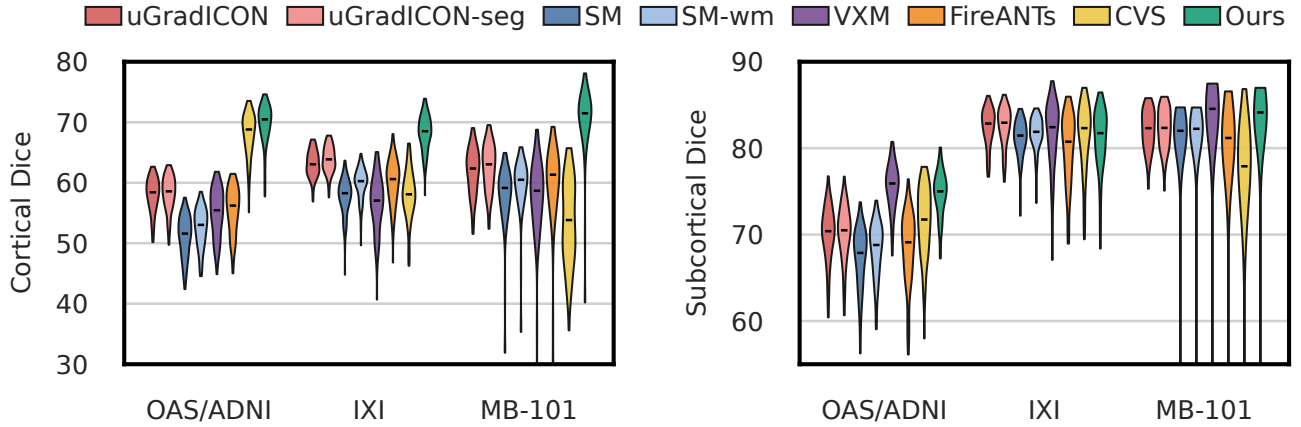


Fig. 2. Violin plot of distributions of Dice score across all subjects in each dataset. Our method (blue) consistently achieves the highest mean Dice for cortical structures, with mass centered near the mean. We preserve subcortical performance while delivering much stronger cortical alignment.

4.2 Results

Table 1 and Figure 2 present results on the 3 datasets. Our method, NeurAlign, consistently achieves the highest cortical Dice score, often by a large margin. For example, it improves by up to 7 Dice points on the held-out Mindboggle dataset. NeurAlign also achieves the highest subcortical Dice score on OASIS-1 & ADNI and Mindboggle. We obtain slightly lower subcortical Dice on IXI (1.5 points at worse), though the difference is not substantial. Across all datasets, NeurAlign achieves statistically better cortical Dice scores ($p < 0.01$) and in many cases a substantial increase (up to 7.5 points over the next best method). We also outperform baselines on whole cerebral cortex alignment (see Suppl. A.2). Finally, NeurAlign produces regular deformation fields with negligible folding. These results demonstrate that we are able to greatly improve cortical structure alignment while maintaining strong performance in substructure alignment and field regularity.

CVS was the slowest method, requiring 2.5 ± 0.5 hours per pair. uGradICON required minutes to perform test-time optimization. NeurAlign and all other baselines converged in milliseconds.

Figure 3 visualizes example registrations and corresponding warps. Our method yields accurate alignments with smooth and regular deformation fields.

4.3 Ablations

To study important aspects of NeurAlign, we conduct two ablation experiments. In the first, we quantify the effect of different model components: label map supervision in training and including the spherical alignment loss. In the second, we quantify the effect of κ , the Dice loss weight parameter.

Results. Table 2 presents ablations of model components, trained with $\kappa = 1.0$ and $\gamma = 0.05$, on the IXI dataset. We train a base VoxelMorph model [15], called Base, and extensions testing components of NeurAlign. We compare the performance when training with supervision of segmentations of subcortical structures,

Base+Dice(subcort), training with segmentations of all structures, Base+Dice(all), training with only the spherical consistency loss, Base+Sphere, and training with all proposed components, Base+Dice(all)+Sphere. We selected $\kappa = 1.0$, as it achieved the best performance for Base+Dice(subcort).

As expected, training with subcortical supervision improved the Dice (subcortical) score compared to the Base model. However, interestingly, training with supervision of all brain structures did not lead to a substantial improvement in the Dice (cortical) score. Only when including our proposed spherical consistency loss (Base+Dice(all)+Sphere) did we observe a substantial increase in the Dice (cortical) score while maintaining the Dice (subcortical) score. Training with just the spherical loss without structural supervision did not improve results, indicating that we also require volumetric structure supervision. This ablation study demonstrates the benefit of our spherical-volumetric consistency constraint in obtain accurate alignment of fine-grained cortical structures.

We assess the impact of κ , the weighting hyperparameter for the Dice loss term (see Suppl. Table A.1, Fig. A.1). As expected, increasing κ improves Dice performance at test time, yielding higher structural overlap. In some regions, however, this comes at the expense of deformation field regularity. Overall, setting $\kappa > 1$ produces models that outperform baselines across most metrics. Consistent with prior work [16, 29], the optimal choice of κ depends on the intended downstream application: for atlas-based segmentation propagation, higher Dice weighting is advantageous, whereas for longitudinal studies, smoother deformation fields may be preferable. Hypernetworks can be considered to select hyperparameters at test-time [29].

5 Discussion

Limitations and Future Work. We presented NeurAlign, a model for unified cortical and subcortical neuroimage registration, evaluated primarily on adult T1w MRI from healthy and Alzheimer’s disease populations. As with any diffeomorphic registration framework, NeurAlign cannot accurately accommodate topology-altering

Table 1. Performance across three datasets. The proposed method (NeurAlign) substantially improves cortical dice and consistently matches or improves subcortical Dice, and produces smooth deformation fields with negligible folding. Mean $\pm$ standard deviation across subjects are indicated.

Dataset	Method	Dice ($\uparrow$) cortical	Dice ($\uparrow$) subcortical	% folds ($\downarrow$)	SD log det J ($\downarrow$)
OASIS-1 & ADNI	uGradICON	0.58 ± 0.028	0.701 ± 0.032	0.647 ± 0.105	0.408 ± 0.127
	uGradICON-seg	0.583 ± 0.028	0.701 ± 0.031	0.688 ± 0.101	0.41 ± 0.126
	SynthMorph	0.511 ± 0.033	0.673 ± 0.034	0.0 ± 0.0	0.197 ± 0.1
	SynthMorph-wm	0.525 ± 0.032	0.682 ± 0.031	0.0 ± 0.0	0.188 ± 0.1
	VoxelMorph	0.551 ± 0.039	0.756 ± 0.026	0.001 ± 0.001	0.479 ± 0.233
	FireANTs	0.554 ± 0.038	0.684 ± 0.041	1.319 ± 0.171	0.5 ± 0.137
	CVS	0.681 ± 0.035	0.715 ± 0.041	1.73 ± 0.321	8.737 ± 3.475
	NeurAlign (Ours)	0.698 ± 0.032	0.747 ± 0.026	0.169 ± 0.037	0.829 ± 0.281
IXI	uGradICON	0.631 ± 0.021	0.824 ± 0.022	0.489 ± 0.1	0.382 ± 0.121
	uGradICON-seg	0.639 ± 0.021	0.825 ± 0.022	0.602 ± 0.109	0.399 ± 0.125
	SynthMorph	0.577 ± 0.03	0.811 ± 0.021	0.0 ± 0.0	0.177 ± 0.069
	SynthMorph-wm	0.599 ± 0.025	0.817 ± 0.019	0.0 ± 0.0	0.167 ± 0.064
	VoxelMorph	0.566 ± 0.045	0.816 ± 0.038	0.001 ± 0.001	0.465 ± 0.17
	FireANTs	0.601 ± 0.037	0.8 ± 0.04	1.382 ± 0.183	0.501 ± 0.133
	CVS	0.582 ± 0.039	0.814 ± 0.038	1.865 ± 0.382	7.591 ± 4.115
	NeurAlign (Ours)	0.683 ± 0.029	0.81 ± 0.036	0.081 ± 0.023	0.712 ± 0.24
Mindboggle	uGradICON	0.618 ± 0.038	0.821 ± 0.023	0.85 ± 0.404	0.436 ± 0.199
	uGradICON-seg	0.626 ± 0.038	0.822 ± 0.023	0.958 ± 0.4	0.445 ± 0.188
	SynthMorph	0.577 ± 0.054	0.806 ± 0.065	0.0 ± 0.0	0.185 ± 0.071
	SynthMorph-wm	0.596 ± 0.048	0.811 ± 0.059	0.0 ± 0.0	0.173 ± 0.062
	VoxelMorph	0.573 ± 0.072	0.827 ± 0.09	0.002 ± 0.002	0.51 ± 0.245
	FireANTs	0.597 ± 0.07	0.79 ± 0.09	1.604 ± 0.347	0.53 ± 0.169
	CVS	0.535 ± 0.07	0.766 ± 0.081	2.164 ± 0.652	7.429 ± 4.654
	NeurAlign (Ours)	0.703 ± 0.06	0.823 ± 0.086	0.174 ± 0.034	0.831 ± 0.261

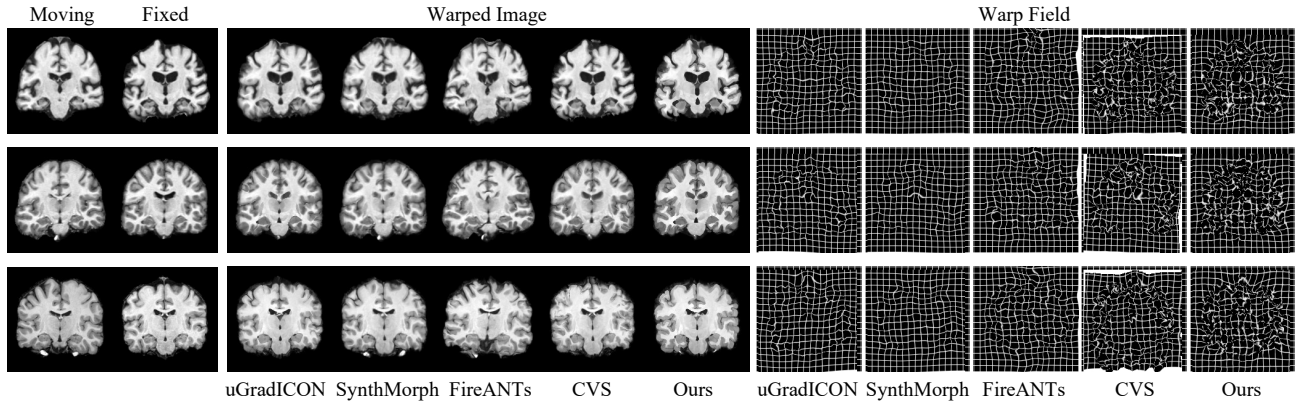


Fig. 3. Example registrations for three image pairs. Columns show the moving and fixed images, followed by warped images produced by uGradICON-seg, SynthMorph-wm, FireANTs, CVS, and our method. Corresponding deformation fields are visualized on the right. Our method (NeurAlign) yields accurate alignments with smooth and regular warp fields.

pathologies. This limitation can be partially mitigated by incorporating pathology masks during training [11]. It also remains unclear how well the model can generalize to lower-quality clinical scans or to pediatric populations. Extending NeurAlign to these more challenging settings is an important direction for future work.

Training our model requires cortical surface extraction and spherical inflation. This preprocessing pipeline may fail on challenging scans. A simple approach for quality control could be to discard failed cases, for example surfaces with many self-intersecting faces,

or to use robust, learning-based extraction methods [10, 30]. In practice, we did not observe any failures on our training data.

Our model was also only tested on the T1w modality. T1w is the most widely used imaging modality for structural imaging studies [9], and it is currently the primary modality from which cortical surfaces can be reliably reconstructed [21, 30]. An interesting future direction is extending NeurAlign to multimodal acquisitions. Additional contrasts such as T2w, FLAIR, or DWI could provide complementary information for subcortical tissue segmentation or characterization. Incorporating these modalities would require multimodal training with all contrasts aligned to the T1w image so that the cortical surfaces remain consistent across channels. Importantly, at inference time, the surfaces would no longer be required, potentially reducing the current dependence on T1w imaging for surface extraction.

A benefit of our formulation is that it can be easily extended to include additional structures with genus-0 topology. For example, this can improve alignment of the hippocampus by providing additional supervision through surface alignment. Further, our consistency loss in Eq. (3) can be adopted for non-spherical surface registration, implicit surfaces, or to perform cortical alignment directly on the surface mesh using frameworks such as Deep Functional Maps [20]. Representations beyond spheres can be used, provided there is a one-to-one map to the cortical mesh.

Conclusion. We introduced NeurAlign, a deformable registration method that achieves accurate cortical and subcortical alignment. NeurAlign parameterizes the deformation field through coupled spherical and volumetric alignment, promoting topological consistency on the cortex.

Our model outperforms state-of-the-art deep learning and foundational models trained or fine-tuned using cortical segmentation label maps. Compared to CVS, the most widely used method in large-scale neuroimaging pipelines, our method achieves substantially better alignment and deformation field regularity on held-out datasets, while running orders of magnitude faster. Further, our method does not require additional inputs (cortical meshes, segmentations) as inputs. This positions our method as a practical alternative to CVS for many neuroimaging applications.

Reproducibility Statement

Our model is summarized in Section 3.3, including descriptions of the 3D and 2D models used and loss functions in training. We provide implementation details including all hyperparameters used,

Table 2. Ablations over training settings on IXI. The proposed spherical alignment consistency loss is necessary to obtain high cortical structure alignment while maintaining strong subcortical alignment.

Ablation	Dice ($\uparrow$) cortical	Dice ($\uparrow$) subcortical	% folds ($\downarrow$)	SD log det J ($\downarrow$)
Base	0.582 $\pm$ 0.045	0.789 $\pm$ 0.049	0.003 $\pm$ 0.003	0.446 $\pm$ 0.172
Base+Dice (subcort)	0.553 $\pm$ 0.045	0.815 $\pm$ 0.038	0.002 $\pm$ 0.001	0.494 $\pm$ 0.185
Base+Dice (all)	0.562 $\pm$ 0.044	0.812 $\pm$ 0.041	0.001 $\pm$ 0.001	0.462 $\pm$ 0.168
Base+Sphere	0.56 $\pm$ 0.03	0.736 $\pm$ 0.047	0.004 $\pm$ 0.002	0.421 $\pm$ 0.152
Base+Dice(all)+Sphere	0.633 $\pm$ 0.028	0.799 $\pm$ 0.038	0.006 $\pm$ 0.003	0.432 $\pm$ 0.137

and data processing details in Section 4.1. Complete code for full reproducibility will be made public upon acceptance.

Acknowledgments

We thank Neel Dey for helpful discussions. Support for this research was provided in part by Quanta Computer Inc., the BRAIN Initiative Cell Atlas Network (BICAN) grants U01MH117023, UM1MH134812 and UM1MH130981, the Brain Initiative Brain Connects consortium (U01NS132181, UM1NS132358-01), the National Institute for Biomedical Imaging and Bioengineering (R01 EB033773, 2R01EB023281, R21EB018907, R01EB019956, P41EB030006), the National Institute on Aging (R21AG082082, 1R01AG064027, R01AG016495, 1R01AG070988), the National Institute of Mental Health (UM1MH130981, R01 MH123195, R01 MH121885, 1RF1MH123195), the National Institute for Neurological Disorders and Stroke, (1U24NS135561-01, R01NS070963, 2R01NS083534, R01NS105820, R25NS125599), the National Institute for Child Health and Human Development (R01HD109436, R00HD101553), and was made possible by the resources provided by Shared Instrumentation Grants 1S10RR023401, 1S10RR019307, and 1S10RR023043. Additional support was provided by the NIH Blueprint for Neuroscience Research (5U01-MH093765), part of the multi-institutional Human Connectome Project. Parts of this work were also supported by the ERC Starting Grant 758800 (EXPROTEA), ERC Consolidator Grant 101087347 (VEGA), as well as gifts from Ansys and Adobe Research. Much of the computation resources required for this research was performed on computational hardware generously provided by the Massachusetts Life Sciences Center (<https://www.masslifesciences.com/>). In addition, BF is an advisor to DeepHealth, a company whose medical pursuits focus on medical imaging and measurement technologies. AD is an advisor to Radence and DeepHealth. BF’s and AD’s interests were reviewed and are managed by Massachusetts General Hospital and Mass. General Brigham in accordance with their conflict of interest policies.

References

- [1] S Mazdak Abulnaga, Oded Stein, Polina Golland, and Justin Solomon. Symmetric volume maps: Order-invariant volumetric mesh correspondence with free boundary. *ACM Transactions on Graphics*, 42(3):1–20, 2023.
- [2] S. Mazdak Abulnaga, Andrew Hoopes, Neel Dey, Malte Hoffmann, Bruce Fischl, John Gutttag, and Adrian Dalca. Multimorph: On-demand atlas construction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 30906–30917, June 2025.
- [3] Sahar Ahmad, Zhengwang Wu, Gang Li, Li Wang, Weili Lin, Pew-Thian Yap, and Dinggang Shen. Surface-constrained volumetric registration for the early developing brain. *Medical Image Analysis*, 58:101540, 2019.
- [4] John Ashburner. A fast diffeomorphic image registration algorithm. *Neuroimage*, 38(1):95–113, 2007.
- [5] 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.
- [6] Brian B Avants, Paul Yushkevich, John Pluta, David Minkoff, Marc Korczykowski, John Detre, and James C Gee. The optimal template effect in hippocampus studies of diseased populations. *Neuroimage*, 49(3):2457–2466, 2010.
- [7] Guha Balakrishnan, Amy Zhao, Mert R Sabuncu, John Gutttag, and Adrian V Dalca. Voxelmorph: a learning framework for deformable medical image registration. *IEEE transactions on medical imaging*, 38(8):1788–1800, 2019.
- [8] Paul J. Besl and Neil D. McKay. A Method for Registration of 3-D Shapes. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1992.
- [9] Gaurav Bhalerao, Grace Gillis, Mohamed Dembele, Sana Suri, Klaus Ebmeier, Johannes Klein, Michele Hu, Clare Mackay, and Ludovica Griffanti. Automated quality control of t1-weighted brain mri scans for clinical research: methods

- comparison and design of a quality prediction classifier. *medRxiv*, pp. 2024–04, 2024.
- [10] Fabian Bongratz, Anne-Marie Rickmann, Sebastian Pölsterl, and Christian Wachinger. Vox2cortex: Fast explicit reconstruction of cortical surfaces from 3d mri scans with geometric deep neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 20773–20783, 2022.
 - [11] Matthew Brett, Alexander P Leff, Chris Rorden, and John Ashburner. Spatial normalization of brain images with focal lesions using cost function masking. *Neuroimage*, 14(2):486–500, 2001.
 - [12] Dongliang Cao, Paul Roetzer, and Florian Bernard. Unsupervised Learning of Robust Spectral Shape Matching. *ACM Transactions on Graphics*, 42(4):132:1–132:15, July 2023.
 - [13] Jieyu Cheng, Adrian V Dalca, Bruce Fischl, Lilla Zöllei, Alzheimer’s Disease Neuroimaging Initiative, et al. Cortical surface registration using unsupervised learning. *NeuroImage*, 221:117161, 2020.
 - [14] Bryan R Conroy, Benjamin D Singer, J Swaroop Guntupalli, Peter J Ramadge, and James V Haxby. Inter-subject alignment of human cortical anatomy using functional connectivity. *NeuroImage*, 81:400–411, 2013.
 - [15] Adrian Dalca, Marianne Rakic, John Guttag, and Mert Sabuncu. Learning conditional deformable templates with convolutional networks. *Advances in neural information processing systems*, 32, 2019.
 - [16] 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.
 - [17] Bob D De Vos, Floris F Berendsen, Max A Viergever, Marius Staring, and Ivana Išgum. End-to-end unsupervised deformable image registration with a convolutional neural network. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings 3*, pp. 204–212. Springer, 2017.
 - [18] Bailin Deng, Yuxin Yao, Roberto M. Dyke, and Juyong Zhang. A Survey of Non-Rigid 3D Registration. *Computer Graphics Forum*, 41(2):559–589, 2022.
 - [19] Christophe Destrieux, Bruce Fischl, Anders Dale, and Eric Halgren. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage*, 53(1):1–15, 2010.
 - [20] Nicolas Donati, Abhishek Sharma, and Maks Ovsjanikov. Deep Geometric Functional Maps: Robust Feature Learning for Shape Correspondence. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8589–8598, Seattle, WA, USA, June 2020. IEEE. ISBN 978-1-72817-168-5.
 - [21] Bruce Fischl. Freesurfer. *Neuroimage*, 62(2):774–781, 2012.
 - [22] Bruce Fischl, Martin I Sereno, and Anders M Dale. Cortical surface-based analysis: Ii: inflation, flattening, and a surface-based coordinate system. *Neuroimage*, 9(2):195–207, 1999.
 - [23] Bruce Fischl, Niranjini Rajendran, Evelina Busa, Jean Augustinack, Oliver Hinds, BT Thomas Yeo, Hartmut Mohlberg, Katrin Amunts, and Karl Zilles. Cortical folding patterns and predicting cytoarchitecture. *Cerebral cortex*, 18(8):1973–1980, 2008.
 - [24] Satrajit S Ghosh, Sita Kakunoori, Jean Augustinack, Alfonso Nieto-Castanon, Ioulia Kovelman, Nadine Gaab, Joanna A Christodoulou, Christina Triantafyllou, John DE Gabrieli, and Bruce Fischl. Evaluating the validity of volume-based and surface-based brain image registration for developmental cognitive neuroscience studies in children 4 to 11 years of age. *Neuroimage*, 53(1):85–93, 2010.
 - [25] Daniel Grzech, Mohammad Farid Azampour, Ben Glocker, Julia Schnabel, Nassir Navab, Bernhard Kainz, and Loïc Le Folgoc. A variational Bayesian method for similarity learning in non-rigid image registration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 119–128, 2022.
 - [26] Malte Hoffmann, Benjamin Billot, Douglas N Greve, Juan Eugenio Iglesias, Bruce Fischl, and Adrian V Dalca. Synthemorph: learning contrast-invariant registration without acquired images. *IEEE transactions on medical imaging*, 41(3):543–558, 2021.
 - [27] Malte Hoffmann, Andrew Hoopes, Bruce Fischl, and Adrian V Dalca. Anatomy-specific acquisition-agnostic affine registration learned from fictitious images. In *Medical Imaging 2023: Image Processing*, volume 12464, pp. 1246402. SPIE, 2023.
 - [28] Malte Hoffmann, Andrew Hoopes, Douglas N Greve, Bruce Fischl, and Adrian V Dalca. Anatomy-aware and acquisition-agnostic joint registration with Synthemorph. *Imaging Neuroscience*, 2:1–33, 2024.
 - [29] Andrew Hoopes, Malte Hoffmann, Bruce Fischl, John Guttag, and Adrian V Dalca. Hypermorph: Amortized hyperparameter learning for image registration. In *International conference on information processing in medical imaging*, pp. 3–17. Springer, 2021.
 - [30] Andrew Hoopes, Juan Eugenio Iglesias, Bruce Fischl, Douglas Greve, and Adrian V Dalca. Topofit: rapid reconstruction of topologically-correct cortical surfaces. *Proceedings of machine learning research*, 172:508, 2022.
 - [31] IXI Consortium. IXI dataset. <https://brain-development.org/ixi-dataset/>. URL <https://brain-development.org/ixi-dataset/>.
 - [32] Rohit Jena, Pratik Chaudhari, and James C Gee. Deep implicit optimization enables robust learnable features for deformable image registration. *Medical Image Analysis*, 103:103577, 2025.
 - [33] Anand Joshi, Richard Leahy, Arthur W Toga, and David Shattuck. A framework for brain registration via simultaneous surface and volume flow. In *International Conference on Information Processing in Medical Imaging*, pp. 576–588. Springer, 2009.
 - [34] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
 - [35] Arno Klein, Satrajit S Ghosh, Forrest S Bao, Joachim Giard, Yrjö Häme, Eliezer Stavsky, Noah Lee, Brian Rossa, Martin Reuter, Elias Chaibub Neto, et al. Mindboggling morphometry of human brains. *PLoS computational biology*, 13(2):e1005350, 2017.
 - [36] Alex D Leow, Igor Yanovsky, Ming-Chang Chiang, Agatha D Lee, Andrea D Klunder, Allen Lu, James T Becker, Simon W Davis, Arthur W Toga, and Paul M Thompson. Statistical properties of jacobian maps and the realization of unbiased large-deformation nonlinear image registration. *IEEE transactions on medical imaging*, 26(6):822–832, 2007.
 - [37] Jian Li, Greta Tuckute, Evelina Fedorenko, Brian L Edlow, Adrian V Dalca, and Bruce Fischl. Josa: Joint surface-based registration and atlas construction of brain geometry and function. *Medical Image Analysis*, 98:103292, 2024.
 - [38] Robin Magnet and Maks Ovsjanikov. Memory-Scalable and Simplified Functional Map Learning. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4041–4050, Seattle, WA, USA, June 2024. IEEE. ISBN 9798350353006.
 - [39] 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.
 - [40] Marc Modat, Gerard R Ridgway, Zeike A Taylor, Manja Lehmann, Josephine Barnes, David J Hawkes, Nick C Fox, and Sébastien Ourselin. Fast free-form deformation using graphics processing units. *Computer methods and programs in biomedicine*, 98(3):278–284, 2010.
 - [41] Tony CW Mok, Zi Li, Yingda Xia, Jiawen Yao, Ling Zhang, Jingren Zhou, and Le Lu. Deformable medical image registration under distribution shifts with neural instance optimization. In *International Workshop on Machine Learning in Medical Imaging*, pp. 126–136. Springer, 2023.
 - [42] Susanne G Mueller, Michael W Weiner, Leon J Thal, Ronald C Petersen, Clifford Jack, William Jagust, John Q Trojanowski, Arthur W Toga, and Laurel Beckett. Ways toward an early diagnosis in alzheimer’s disease: the alzheimer’s disease neuroimaging initiative (adni). *Alzheimer’s & dementia*, 1(1):55–66, 2005.
 - [43] Maks Ovsjanikov, Mirela Ben-Chen, Justin Solomon, Adrian Butscher, and Leonidas Guibas. Functional maps: A flexible representation of maps between shapes. *ACM Transactions on Graphics*, 31(4):1–11, August 2012.
 - [44] Gheorghe Postelnicu, Lilla Zöllei, and Bruce Fischl. Combined volumetric and surface registration. *IEEE transactions on medical imaging*, 28(4):508–522, 2008.
 - [45] Marianne Rakic, Andrew Hoopes, Mazdak S. Abulnaga, Mert R. Sabuncu, John V. Guttag, and Adrian V. Dalca. Atlasmorph: Learning conditional deformable templates for brain mri. *Medical Image Analysis*, 110:103893, 2026. ISSN 1361-8415.
 - [46] Marc-Michel Rohé, Manasi Datar, Tobias Heimann, Maxime Sermesant, and Xavier Pennec. Svf-net: learning deformable image registration using shape matching. In *Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I 20*, pp. 266–274. Springer, 2017.
 - [47] Karl Rohr, H Siegfried Stieh, Rainer Sprengel, Thorsten M Buzug, Jürgen Weese, and MH Kuhn. Landmark-based elastic registration using approximating thin-plate splines. *IEEE Transactions on medical imaging*, 20(6):526–534, 2001.
 - [48] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention-MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pp. 234–241. Springer, 2015.
 - [49] Daniel Rueckert, Luke I Sonoda, Carmel Hayes, Derek LG Hill, Martin O Leach, and David J Hawkes. Nonrigid registration using free-form deformations: application to breast mr images. *IEEE transactions on medical imaging*, 18(8):712–721, 1999.
 - [50] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *Proceedings Third International Conference on 3-D Digital Imaging and Modeling*, pp. 145–152, May 2001.
 - [51] Florent Ségonne, Jenni Pacheco, and Bruce Fischl. Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. *IEEE transactions on medical imaging*, 26(4):518–529, 2007.
 - [52] Nicholas Sharp, Souhaib Attaiki, Keenan Crane, and Maks Ovsjanikov. Diffusion-Net: Discretization Agnostic Learning on Surfaces. *ACM Transactions on Graphics*, 41(3):1–16, June 2022.

- [53] Hessam Sokooti, Bob De Vos, Floris Berendsen, Boudewijn PF Lelieveldt, Ivana Išgum, and Marius Staring. Nonrigid image registration using multi-scale 3d convolutional neural networks. In *Medical Image Computing and Computer Assisted Intervention - MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I 20*, pp. 232–239. Springer, 2017.
- [54] Olga Sorkine and Marc Alexa. As-rigid-as-possible surface modeling. In *Proceedings of the Fifth Eurographics Symposium on Geometry Processing, SGP '07*, pp. 109–116, Goslar, DEU, July 2007. Eurographics Association. ISBN 978-3-905673-46-3.
- [55] J. (Jean) Talairach and Pierre Tournoux. Co-planar stereotaxic atlas of the human brain—3-dimensional proportional system: An approach to cerebral imaging. *Clinical Radiology*, 41:219, 1990.
- [56] Paul M Thompson, Kiralee M Hayashi, Rebecca A Dutton, MING-CHANG CHI-ANG, Alex D Leow, Elizabeth R Sowell, Greig De Zubicaray, James T Becker, Oscar L Lopez, Howard J Aizenstein, et al. Tracking alzheimer’s disease. *Annals of the New York Academy of Sciences*, 1097(1):183–214, 2007.
- [57] Lin Tian, Hastings Greer, Roland Kwitt, Francois-Xavier Vialard, Raúl San José Estépar, Sylvain Bouix, Richard Rushmore, and Marc Niethammer. unigradicon: A foundation model for medical image registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 749–760. Springer, 2024.
- [58] Alan Tucholka, Virgile Fritsch, Jean-Baptiste Poline, and Bertrand Thirion. An empirical comparison of surface-based and volume-based group studies in neuroimaging. *Neuroimage*, 63(3):1443–1453, 2012.
- [59] Clarisse Weischedel. *A Discrete Geometric View on Shear-Deformable Shell Models*. PhD thesis, Georg-August-Universität, Göttingen, August 2012.
- [60] Xiao Yang, Roland Kwitt, Martin Styner, and Marc Niethammer. Quicksilver: Fast predictive image registration—a deep learning approach. *NeuroImage*, 158: 378–396, 2017.
- [61] BT Thomas Yeo, Mert R Sabuncu, Tom Vercauteren, Nicholas Ayache, Bruce Fischl, and Polina Golland. Spherical demons: fast diffeomorphic landmark-free surface registration. *IEEE transactions on medical imaging*, 29(3):650–668, 2009.
- [62] Fenqiang Zhao, Shunren Xia, Zhengwang Wu, Dingna Duan, Li Wang, Weili Lin, John H Gilmore, Dinggang Shen, and Gang Li. Spherical u-net on cortical surfaces: methods and applications. In *Information Processing in Medical Imaging: 26th International Conference, IPMI 2019, Hong Kong, China, June 2–7, 2019, Proceedings 26*, pp. 855–866. Springer, 2019.
- [63] Fenqiang Zhao, Zhengwang Wu, Li Wang, Weili Lin, John H Gilmore, Shunren Xia, Dinggang Shen, and Gang Li. Spherical deformable u-net: Application to cortical surface parcellation and development prediction. *IEEE transactions on medical imaging*, 40(4):1217–1228, 2021.
- [64] Lilla Zöllei, Allison Stevens, Kristen Huber, Sita Kakunoori, and Bruce Fischl. Improved tractography alignment using combined volumetric and surface registration. *Neuroimage*, 51(1):206–213, 2010.

A Supplementary Material

A.1 Ablation over Dice Loss Weight

In this section, we conduct ablations over the Dice loss weight κ . We use the IXI dataset and a sphere loss weight $\gamma = 0.05$. As demonstrated on Table 2, the inclusion of the soft Dice loss on cortical and subcortical structures in combination with our proposed spherical alignment loss is necessary to produce accurate cortical and subcortical alignments. However, this comes at the cost of field regularity. Table A.1 presents the results. Consistent increases in Dice (cortical) are observed with increasing κ , but Dice (subcortical) starts to saturate for $\kappa > 5$. This comes at the expense of field regularity, as measured by the % of folds and SDlog det J .

Figure A.1 presents example images and warps varying κ . As expected, deformation fields become more irregular with increasing κ , though presented ranges are reasonable. The optimal choice of κ depends on the intended downstream application. The user has control over the desired strength of structural alignment over field regularity and smoothness. We note however that our model produces accurate matchings with regular fields for a large range of κ values.

A.2 Whole Cerebral Cortex Alignment Performance

Our evaluation on Table 1 and Fig. 2 focused on quantifying performance of aligning cortical parcellations within the cerebral cortex. We hypothesized that the spherical alignment signal will be most important for aligning the homologous folds as the image intensity information is unable to capture the functional regions within the cortex. To test this hypothesis, we compared Dice score in aligning cortical parcellations (Dice cortical) with Dice score aligning the entire cerebral cortex (Dice CC), which groups all parcellations as one class. Table A.2 presents the results.

On two of the three datasets, our model achieves the highest Dice (CC) performance, while CVS outperforms our method by 1.2 Dice points on the OASIS-1 & ADNI test set. Interestingly, all baseline methods show improved performance in aligning the cerebral cortex and differences across baselines are smaller. This indicates that purely volumetric approaches may be sufficient at capturing the whole cerebral cortex structure. However, when trying to align the cortical parcellations, the benefit of our formulation is immediate, as we obtain 1.9 – 7.5 point Dice increases over the next best method across all datasets. The purely volumetric approaches show substantially lower performance than our method, as well.

Table A.1. Ablations over Dice loss weight κ on the IXI dataset.

Dice loss weight	Dice ($\uparrow$) cortical	Dice ($\uparrow$) subcortical	% folds ($\downarrow$)	SD log det J ($\downarrow$)
$\kappa = 0.5$	0.613 ± 0.029	0.787 ± 0.041	0.014 ± 0.005	0.459 ± 0.15
$\kappa = 1.0$	0.633 ± 0.028	0.799 ± 0.038	0.006 ± 0.003	0.432 ± 0.137
$\kappa = 2.0$	0.637 ± 0.03	0.793 ± 0.037	0.016 ± 0.006	0.462 ± 0.141
$\kappa = 5$	0.653 ± 0.029	0.796 ± 0.039	0.046 ± 0.015	0.597 ± 0.188
$\kappa = 10$	0.683 ± 0.029	0.81 ± 0.036	0.081 ± 0.023	0.712 ± 0.24
$\kappa = 50$	0.691 ± 0.028	0.799 ± 0.038	0.848 ± 0.135	1.294 ± 0.414

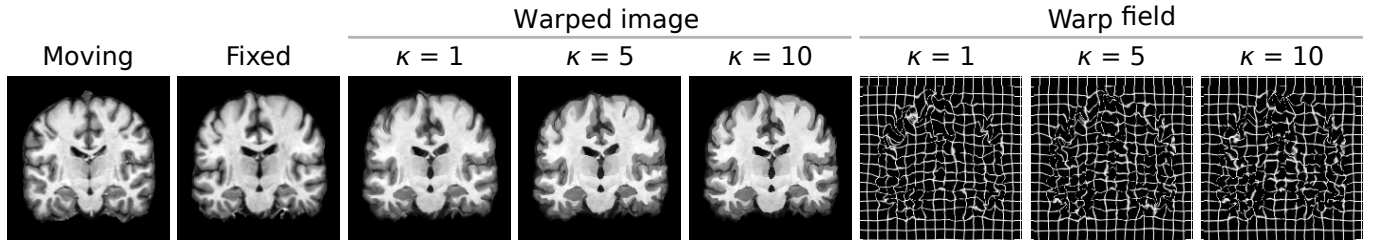


Fig. A.1. Example registrations for one image while varying κ . Columns show the moving and fixed images, followed by warped images and corresponding deformation fields are visualized on the right. Deformation fields become more irregular with increasing κ , however all demonstrated values produce accurate and smooth alignments.

Table A.2. Performance of all methods across three datasets, focusing on structural alignment on the graymatter. Dice (CC) indicates the dice score in aligning the entire cerebral cortex, while Dice cortical refers to the parcellated substructures within the cerebral cortex. The benefit of our method in aligning substructures within the cortex (Dice cortical) is immediately apparent, as we consistently achieve the highest score, including up to a 7.5 dice score improvement over the next method. When considering the cerebral cortex as one structure, performance differences decrease, though our method still performs the strongest on 2 of the 3 datasets. The benefit of the spherical registration is immediately apparent as it is necessary to align functional regions in the cortex where image-only approaches fail. Mean $\pm$ standard deviation across subjects are indicated.

Dataset	Method	Dice (CC) ($\uparrow$) left	Dice (CC) ($\uparrow$) right	Dice ($\uparrow$) cortical	% folds ($\downarrow$)
OASIS-1 & ADNI	uGradICON	0.711 ± 0.029	0.71 ± 0.026	0.58 ± 0.028	0.647 ± 0.105
	uGradICON-seg	0.712 ± 0.03	0.711 ± 0.027	0.583 ± 0.028	0.688 ± 0.101
	SynthMorph	0.626 ± 0.035	0.628 ± 0.032	0.511 ± 0.033	0.0 ± 0.0
	SynthMorph-wm	0.617 ± 0.035	0.62 ± 0.031	0.525 ± 0.032	0.0 ± 0.0
	VoxelMorph	0.72 ± 0.031	0.72 ± 0.03	0.551 ± 0.039	0.001 ± 0.001
	FireANTs	0.703 ± 0.034	0.703 ± 0.033	0.554 ± 0.038	1.319 ± 0.171
	CVS	0.777 ± 0.03	0.775 ± 0.03	0.681 ± 0.035	1.73 ± 0.321
	NeurAlign (Ours)	0.765 ± 0.031	0.764 ± 0.025	0.698 ± 0.032	0.169 ± 0.037
IXI	uGradICON	0.74 ± 0.02	0.746 ± 0.017	0.631 ± 0.021	0.489 ± 0.1
	uGradICON-seg	0.749 ± 0.02	0.755 ± 0.017	0.639 ± 0.021	0.602 ± 0.109
	SynthMorph	0.691 ± 0.02	0.693 ± 0.02	0.577 ± 0.03	0.0 ± 0.0
	SynthMorph-wm	0.683 ± 0.021	0.684 ± 0.021	0.599 ± 0.025	0.0 ± 0.0
	VoxelMorph	0.736 ± 0.033	0.737 ± 0.032	0.566 ± 0.045	0.001 ± 0.001
	FireANTs	0.753 ± 0.025	0.754 ± 0.024	0.601 ± 0.037	1.382 ± 0.183
	CVS	0.748 ± 0.041	0.754 ± 0.036	0.582 ± 0.039	1.865 ± 0.382
	NeurAlign (Ours)	0.758 ± 0.023	0.758 ± 0.021	0.683 ± 0.029	0.081 ± 0.023
Mindboggle	uGradICON	0.722 ± 0.051	0.731 ± 0.044	0.618 ± 0.038	0.85 ± 0.404
	uGradICON-seg	0.73 ± 0.049	0.738 ± 0.045	0.626 ± 0.038	0.958 ± 0.4
	SynthMorph	0.697 ± 0.03	0.701 ± 0.036	0.577 ± 0.054	0.0 ± 0.0
	SynthMorph-wm	0.685 ± 0.031	0.688 ± 0.038	0.596 ± 0.048	0.0 ± 0.0
	VoxelMorph	0.736 ± 0.055	0.744 ± 0.055	0.573 ± 0.072	0.002 ± 0.002
	FireANTs	0.742 ± 0.05	0.748 ± 0.051	0.597 ± 0.07	1.604 ± 0.347
	CVS	0.689 ± 0.091	0.693 ± 0.088	0.535 ± 0.07	2.164 ± 0.652
	NeurAlign (Ours)	0.782 ± 0.04	0.785 ± 0.041	0.703 ± 0.06	0.173 ± 0.034