



Diffeomorphic Registration With Intensity Transformation and Missing Data: Application to 3D Digital Pathology of Alzheimer's Disease

Daniel Tward^{1,2*}, Timothy Brown², Yusuke Kageyama^{3†}, Jaymin Patel¹, Zhipeng Hou⁴, Susumu Mori⁴, Marilyn Albert⁵, Juan Troncoso³ and Michael Miller^{1,2}

¹ Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, United States, ² Center for Imaging Science, Johns Hopkins University, Baltimore, MD, United States, ³ Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD, United States, ⁴ Department of Radiology, Johns Hopkins University School of Medicine, Baltimore, MD, United States, ⁵ Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, United States

This paper examines the problem of diffeomorphic image registration in the presence of differing image intensity profiles and sparsely sampled, missing, or damaged tissue. Our motivation comes from the problem of aligning 3D brain MRI with 100-micron isotropic resolution to histology sections at $1 \times 1 \times 1,000$ -micron resolution with multiple varying stains. We pose registration as a penalized Bayesian estimation, exploiting statistical models of image formation where the target images are modeled as sparse and noisy observations of the atlas. In this injective setting, there is no assumption of symmetry between atlas and target. Cross-modality image matching is achieved by jointly estimating polynomial transformations of the atlas intensity. Missing data is accommodated via a multiple atlas selection procedure where several atlas images may be of homogeneous intensity and correspond to "background" or "artifact." The two concepts are combined within an Expectation-Maximization algorithm, where atlas selection posteriors and deformation parameters are updated iteratively and polynomial coefficients are computed in closed form. We validate our method with simulated images, examples from neuropathology, and a standard benchmarking dataset. Finally, we apply it to reconstructing digital pathology and MRI in standard atlas coordinates. By using a standard convolutional neural network to detect tau tangles in histology slices, this registration method enabled us to quantify the 3D density distribution of tauopathy throughout the medial temporal lobe of an Alzheimer's disease postmortem specimen.

Keywords: neuroimaging, digital pathology, histology, brain mapping, image registration, missing data

1. INTRODUCTION

High-throughput neuroinformatics and image analysis are emerging in neuroscience (Miller et al., 2013a; Mori et al., 2013). Atlas-based image analysis plays a key role, as it enables information encoded by millions of independent voxel measurements to be reconstructed in ontologies of the roughly 100 evolutionarily stable structures. At the 1-millimeter scale, there are many atlases,

OPEN ACCESS

Edited by:

John Ashburner, University College London, United Kingdom

Reviewed by:

Suyash P. Awate, Indian Institute of Technology Bombay, India Andrzej Skalski, AGH University of Science and Technology, Poland Marek Wodzinski, AGH University of Science and Technology, Poland, in collaboration with reviewer AS

> *Correspondence: Daniel Tward dtward@cis.jhu.edu

[†]Present address:

Yusuke Kageyama, Molecular Neuroscience Research Center, Shiga University of Medical Science, Otsu, Japan

Specialty section:

This article was submitted to Brain Imaging Methods, a section of the journal Frontiers in Neuroscience

Received: 13 June 2019 Accepted: 14 January 2020 Published: 11 February 2020

Citation:

Tward D, Brown T, Kageyama Y, Patel J, Hou Z, Mori S, Albert M, Troncoso J and Miller M (2020) Diffeomorphic Registration With Intensity Transformation and Missing Data: Application to 3D Digital Pathology of Alzheimer's Disease. Front. Neurosci. 14:52. doi: 10.3389/fnins.2020.00052

1

including Tailarach coordinates (Talairach and Szikla, 1980) and the Montreal Neurological Institute (MNI) (Evans et al., 1993) and Mori's diffusion tensor imaging (DTI) white matter (Mori et al., 2005) atlases, which define the locations of neuroanatomical structures as well as important structural and functional properties such as volume, shape, blood oxygen leveldependent (BOLD) signals, etc. At the millimeter scale, there have been several approaches for mapping onto atlases (Miller et al., 1993; Grenander and Miller, 1998; Toga and Thompson, 2001; Thompson and Toga, 2002; Ashburner and Friston, 2007; Ashburner, 2009), atlas estimation (Durrleman et al., 2008; Vialard et al., 2011), and applications in white matter (Zhang et al., 2010b) or even cardiac imaging (Zhang et al., 2010a; Ardekani et al., 2012). At micron and meso-scales, there are several atlases including Mori's and the Allen brain atlas (Chuang et al., 2011; Sunkin et al., 2013) with their associated region and cell-types. Many of the brain mapping algorithms have been extended to micron scales, such as for CLARITY (Chung et al., 2013; Epp et al., 2015; Kutten et al., 2017).

However, many of the dense brain mapping methods have been based on high-quality datasets, in which image collection is highly controlled and similar image modalities are being registered. In the work described here, we focus on an application that is ubiquitous in digital pathology, where micron-thick tissue slices are prone to damage and are sparse, implying large numbers of missing sections. As well, many stains are often used, which results in a multitude of contrast variations between imaged sections. In the work proposed here, we develop a generative probabilistic model that accounts for differences in shape, contrast, and sparsity associated with censoring of data samples using the random orbit model of Computational Anatomy (Grenander and Miller, 1998, 2007), in which the space of histological images is an injection into the orbit of exemplar templates under both smooth diffeomorphic coordinate transformation and image contrast transformation. The models for coordinate transformations are taken from diffeomorphometry (Miller et al., 2014, 2018). We pose registration as a penalized Bayesian estimation, exploiting statistical models of image formation where the target images are modeled as sparse and noisy observations of the atlas. The correspondence between histology images and 3D atlases is an injection, not an invertible diffeomorphism. The penalized Bayes estimator we derive applies transformations only to the atlas, removing the assumption of symmetry between atlas and target. We accommodate the space of differing contrasts associated with different histological stainings (such as tau, amyloid, myelin, Nissl, and fluorescence) by modeling the space of contrasts via polynomial functions of the atlas. Application of higher-order polynomials describes non-monotone transformations that swap the order of intensities, with first-order monotone polynomials reducing to affine transformations given by the normalized crosscorrelation cost function. To accommodate effects such as folding and distortion, we include additional homogeneous atlases and model each pixel in an observed image as a realization of one transformed atlas from this family. The atlas label at each pixel is interpreted as missing data, with the conditional mean computed

to estimate them using the Expectation-Maximization algorithm (Dempster et al., 1977).

Alternatives to the approach presented here for missing data have included masking of the image similarity objective function (Brett et al., 2001; Stefanescu et al., 2004) or filling with a specific image intensity or texture (Sdika and Pelletier, 2009), as well as explicit joint estimation for missing data, such as that associated with excised tissue (Nithiananthan et al., 2012) or occlusion by lesions (Yoo et al., 2002; Avants et al., 2014), tuberculosis (Vidal et al., 2009), or tumor (Zacharaki et al., 2008).

Registration between different modalities in the presence of missing data is more challenging because image abnormalities are difficult to detect when their expected contrast is unknown. Our approach of attempting to solve the Bayes problem, removing the nuisance variables as much as possible via the Expectation-Maximization algorithm, is more similar to (Periaswamy and Farid, 2006; Chitphakdithai and Duncan, 2010a). Objective functions for multimodality image similarity, such as normalized cross-correlation (Avants et al., 2006; Wu and Tang, 2018), mutual information (Mattes et al., 2003; Pluim et al., 2003), or local structural information (Heinrich et al., 2012; Wachinger and Navab, 2012; Bashiri et al., 2019) do not easily extend to an expectation maximization setting because they do not correspond to a data log-likelihood. Our framework applies to general image registration problems and differs from the methods surveyed above in several important respects. The method proposed here supports large deformations within the computational anatomy random orbit model, unlike (Periaswamy and Farid, 2006; Bashiri et al., 2019), which considers low-dimensional scale rotation and translation dimensions. Secondly, we introduce the many dimensions of polynomials to overcome the nonmonotonicity of the transformation for crossing modalities for histology jointly with deformation, unlike example (Periaswamy and Farid, 2006), which considers multimodality registration and missing tissue separately. Finally, by introducing the EM algorithm, we approximate the full Bayes problem of removing the nuisance dimensions of missing and/or distorted tissue, thereby in part benefiting from the reduced dimensional setting of EM as a solution. The potential impact of these methods has been recognized by the community, and a recent review by Pichat et al. (2018) summarizes work in this area over the past several decades. Artifacts addressed by our method (folds, tears, cracks, and holes) are specifically described as challenges but are not modeled by the registration approaches surveyed. In the survey, these factors are handled by techniques such as manual labeling of artifact regions or modified image similarity metrics such as sum of absolute deviation that avoid strongly penalizing outliers.

We apply these techniques to an important application in Alzheimer's disease (AD), computing the 3D density of tau neurofibrillary tangles, a key pathologic feature of AD (Mirra et al., 1997). Since Braak and Braak's original 1991 staging of Alzheimer's (Braak et al., 2006), it has become clear that one of the earliest locations of degeneration associated with tauopathy is the lateral boundary of the entorhinal cortex, subsequently spreading throughout the medial temporal lobe. In our previous works

(Miller et al., 2013b, 2015; Younes et al., 2014, 2019; Tward et al., 2017; Kulason et al., 2019) based on 1.5 T Magnetic Resonance Imaging (MRI), we demonstrated that there is evidence of consistent atrophy, as manifested by thickness and volume measures. This appears earliest in the entorhinal cortex, with subsequent statistically significant changes seen in the amygdala and hippocampus. Aligning histological sections (stained for tau, amyloid, and myelin) with the dense 3D coordinates of individual brains offers the opportunity to associate the histological changes to pathological tau observed post mortem with structural changes observed clinically. We detect tau tangles using a standard deep convolutional neural network, where 56 \times 56 pixel patches are used to classify the center of the patch. The network is applied to the whole image through overlapping patches in a scanning window approach, shifting by 1 pixel at a time. As this is a largescale task but not a particularly challenging one, this architecture was taken from an official TensorFlow tutorial, and our work here is not intended to advance the state of the art in neural network design.

At the time of manuscript preparation, code to perform 3D registration with this algorithm is available on GitHub (https://github.com/dtward/image_lddmm_tensorflow), implemented in Python¹ with high-performance computing handled using TensorFlow² and available on any operating system that runs Python and TensorFlow. This repository includes example usage in the form of Jupyter notebooks, with figures and explanations of each step, for several datasets covering human, mouse, and rat. The implementation of our algorithm lddmm.py includes descriptions of each step in the form of comments. Our most up-to-date registration code will be described at³.

Our central contribution is to build a registration algorithm by extending the computational anatomy random orbit model in two ways: first, by including nonmonotonic contrast variation in addition to geometric variation, and second, by simultaneously accommodating artifacts or missing tissue through an EM algorithm. Our paper is structured as follows. In section 2, we describe our registration algorithm in three steps: (i) reviewing a traditional approach, (ii) adding contrast variation, and (iii) adding missing data. In section 3, we describe five experiments used to illustrate and validate our method: (i) with a simulated image, (ii) examining our first and second contribution in isolation and together using histology data, (iii) quantitatively demonstrating that our accuracy performance is similar to stateof-the-art alternatives using the CIMA benchmarking dataset described in Borovec et al. (2018), which includes contrast variation but not missing tissue, (iv) quantifying our accuracy via Dice overlap in a missing data setting, and (v) applying our method to Alzheimer's disease research, detecting tau tangles in 2D histology and registering them to the standard 3D coordinate system of the Mai Paxinos Voss atlas (Mai and Paxinos, 2011). In section 4, we present the results of these five experiments. Finally, in section 5, we discuss the implications of this work.

¹https://www.python.org

2. METHODS

In this section, we first review the optimality conditions for the original LDDMM image registration algorithm. We then show our first contribution, how they are modified to include contrast variation, and our second contribution, how they are modified to simultaneously include missing data. In each case, we present three sets of equations: (i) a generative model, (ii) an objective function to be minimized, and (iii) optimality conditions. We conclude this section by summarizing our registration algorithm.

2.1. Background: LDDMM and the Geometric Transformation Problem

Our generative model for the space of observed images builds upon the deformable templates of (Grenander, 1994; Grenander and Miller, 2007). In the original setting, observed images are spatial transformations of grayscale atlases $I: X \subset \mathbb{R}^3 \to \mathbb{R}$.

Geometric differences are modeled as diffeomorphisms, $\varphi \in Diff: X \to X$, which are generated from flows of smooth velocity fields $v_t, t \in [0, 1]$ as in (Grenander and Miller, 1998).

$$\dot{\phi}_t = v_t(\phi_t), \quad \phi_0 = id, \quad \varphi = \phi_1,$$
 (1)

Observed images J are modeled as conditionally Gaussian random fields, with the mean given by the deformed template, and the addition of white noise:

$$J(\cdot) = I(\varphi^{-1}(\cdot)) + \text{noise}(\cdot)$$

With noise variance σ_M^2 , this gives a log-likelihood function only of the parameters

$$-\ell(J;\varphi) = rac{1}{2\sigma_M^2} \|J - I \circ \varphi^{-1}\|_{L_2}^2$$

Because velocity fields are not finite-dimensional, they are regularized by introducing a Sobolev norm as a running penalty, weighted by the parameter σ_R^2 , giving an objective function

$$E(\varphi) = \frac{1}{2\sigma^2} \int_0^1 \int_X Av_t \cdot v_t dx dt + \frac{1}{2\sigma_M^2} \|J - I \circ \varphi^{-1}\|_{L_2}^2$$

Here, *A* is a differential operator of the form $(id - a^2 \Delta)^4$, where *a* is a smoothness length scale.

This problem was solved in Beg et al. (2005) with necessary conditions

$$\nabla E = \frac{1}{\sigma_R^2} A v_t - \frac{1}{\sigma_M^2} \nabla (I(\varphi_t^{-1})) (I(\varphi_t^{-1}) - J(\varphi_{1t}^{-1})) |D\varphi_{1t}^{-1}| = 0$$
(2)

where $\varphi_{1t} = \varphi_t \circ \varphi_1^{-1}$ and $|D\varphi|$ is determinant of the Jacobian matrix.

²https://www.tensorflow.org

³https://neurodata.io/reg

2.2. Polynomial, Non-monotonic Mappings, and the Contrast Transformation Problem

Our first contribution is to expand the generative model to allow observed images to have a different non-monotonically transformed contrast profile from the atlas. We consider atlases $I: X \to \mathbb{R}^N$ and observed images $J: X \to \mathbb{R}^M$. These may be single-valued (N = 1), such as a T1 MRI, or multi-valued, such as red-green-blue (RGB, N = 3). We model this change in contrast by a polynomial function of specified degree with coefficients θ , denoted F_{θ} .

The generative model describes J as a conditionally Gaussian M vector field with each component of noise independent,

$$J(\cdot) = F_{\theta}[I(\varphi^{-1}(\cdot))] + \text{noise}(\cdot).$$

Our objective function is

$$E(\varphi,\theta) = \frac{1}{2\sigma_R^2} \int_0^1 \int_X Av_t \cdot v_t dx dt + \frac{1}{2\sigma_M^2} \|J - F_\theta (I \circ \varphi^{-1})\|_{L_2}^2$$
(3)

where the L_2 norm sums over each component of vectors in \mathbb{R}^M . Our two necessary conditions are

$$\nabla_{\nu}E = \frac{1}{\sigma_{R}^{2}}A\nu_{t} - \frac{1}{\sigma_{M}^{2}}\nablaF_{\theta}(I(\varphi_{t}^{-1})) - J(\varphi_{1t}^{-1}))|D\varphi_{1t}^{-1}| = 0$$
(4a)

$$\nabla_{\theta} E = \int_{X} F_{\theta}(I \circ \varphi_{1}^{-1}(x)) \cdot \frac{d}{d\theta} F_{\theta}(I \circ \varphi_{1}^{-1}(x)) dx$$
$$- \int_{X} J(x) \cdot \frac{d}{d\theta} F_{\theta}(I \circ \varphi_{1}^{-1}(x)) dx = 0$$
(4b)

The first equation (Equation 4a) is a minor modification to Beg's algorithm (2), and the second (Equation 4b) solves for θ as the solution to a least-squares problem. Generally, F_{θ} is nonmonotic polynomial, but when it is an affine transformation, this algorithm is equivalent to minimizing a normalized cross-correlation loss function (see **Appendix 1** in Supplementary Material).

2.3. EM Algorithm and the Missing Data Problem

Our second contribution is to expand the generative model to allow observed images to have a different identity at each pixel. For example, pixels showing torn or folded tissue may be labeled as "background" or "artifact" (respectively) and not correspond to any location in the atlas through the transformation φ . In this setting, missing data takes the form of per-pixel labels that indicate which member of a family of atlases the intensity at that pixel comes from. Thus, the missing tissue (due to a tear or a fold) would be modeled by a "background" atlas, and corrupted pixels due to a streak or a smudge would be modeled by a "artifact" atlas. We approach the problem of missing or censored data using the Expectation-Maximization algorithm (Dempster et al., 1977). Our new generative model describes J with multiple possible means and variances at each pixel i

$$J_{i} = \begin{cases} F_{\theta}[I(\varphi^{-1})]_{i} + \text{noise}_{i}^{M} & \text{if voxel } i \text{ is tissue} \\ \mu_{A} + \text{noise}_{i}^{A} & \text{if voxel } i \text{ is artifact} \\ \mu_{B} + \text{noise}_{i}^{B} & \text{if voxel } i \text{ is background} \end{cases}$$
(5)

the noise components have variance σ_M^2 , σ_A^2 , and σ_B^2 .

Our EM algorithm for jointly estimating the deformation, contrast transformation, and class of each voxel is derived in **Appendix 2** in Supplementary Material. At each M step, we update our transformation parameters by minimizing the objective function

$$E(\varphi,\theta) = \frac{1}{2\sigma_R^2} \int_0^1 \int_X Av_t \cdot v_t dx dt + \frac{1}{2\sigma_M^2} \sum_i |J_i - F_\theta [I \circ \varphi^{-1}]_i)|^2 \pi_i^M$$
(6)

where π_i^M is the posterior probability that voxel *i* corresponds to our atlas image (rather than background or artifact). Our necessary conditions are

$$\nabla_{\nu}E = \frac{1}{\sigma_{R}^{2}}A\nu_{t} - \frac{1}{\sigma_{M}^{2}}\nablaF_{\theta}(I(\varphi_{t}^{-1})) - J(\varphi_{1t}^{-1}))|D\varphi_{1t}^{-1}|\pi^{A} \circ \varphi_{1t}^{-1} = 0$$
(7a)

$$\nabla_{\theta} E = \sum_{i} F_{\theta} [I \circ \varphi_{1}^{-1}(x)]_{i} \cdot \frac{1}{d\theta} F_{\theta} [I \circ \varphi_{1}^{-1}(x)]_{i} \pi_{i}^{M}$$
$$- \sum_{i} J_{i} \cdot \frac{d}{d\theta} F_{\theta} [I \circ \varphi_{1}^{-1}(x)]_{i} \pi_{i}^{M} = 0$$
(7b)

$$\mu_A = \sum_i J_i \pi_i^A / \sum_i \pi_i^A, \qquad \mu_B = \sum_i J_i \pi_i^B / \sum_i \pi_i^B \quad (7c)$$

Here, π^M is interpreted as a continuous function using linear interpolation, and π_i^A, π_i^B are the posterior probability that voxel *i* corresponds to an artifact or background (respectively). The first equation (Equation 7a) differs from Equation (4a) only by the factor π^M , and the second equation (Equation 7b) differs from Equation (4b) by the factor π^M . For technical reasons, we use discrete notation in Equations (7a) and (7b).

2.4. Optimization Algorithm

Our resulting EM-LDDMM optimization algorithm is summarized in Algorithm 1. See the **Appendix 2** in Supplementary Material for a proof that this is an EM algorithm and thus monotonically increasing in likelihood. **Algorithm 1:** EM-LDDMM for image registration with intensity transformations and missing data. Below, we use two possibilities for missing data, artifact, or background.

Input data:

Atlas image *I* and target image *J*.

Input parameters:

 $\sigma_M, \sigma_A, \sigma_B$ (weights in cost function),

a (length scale for transformation smoothness),

O (order of polynomial for contrast mapping). **Initialization**:

 $v_t = 0$ (zero velocity field for identity transform), $\theta = (0, 1, ...)$ polynomial coefficients for identity transform,

 μ_A, μ_B artifact and background intensities.

Optimization: Repeat until convergence.

1. E step:

Compute π_a for matching, artifact, and background using Equation (9).

- 2. M step: Repeat until convergence.
 - (a) Forward pass: Compute φ_t^{-1} from v_t for $t \in [0, 1]$ using Equation (1), and compute $F_{\theta}[I(\varphi_1^{-1})]$.
 - (b) **Objective:** Evaluate the objective function for v_t , $F_{\theta}[I(\varphi_1^{-1})]$ and *J* using Equation (6).
 - (c) **Backward pass**: Compute θ , μ_A , μ_B by weighted leastsquares using Equation (7b) and Equation (7c). Compute φ_t from v_t using Equation (1) and update v_t by gradient descent using Equation (7a).

Output: Forward and inverse transforms $\varphi_1, \varphi_1^{-1}$ (optionally output other parameters of interest).

3. EXPERIMENTAL SETTING AND MATERIAL

In this section, we describe the imaging data acquired as part of our Alzheimer's disease study and the five registration experiments used to illustrate and validate our method.

3.1. Post-mortem Imaging

Preparation and scanning of brain tissue were performed by the neuropathological team at the Johns Hopkins Brain Resource Center (BRC) and the laboratory of Dr. Susumu Mori. The specimen was a 1,290 g brain from a 93-year-old male, with a clinical diagnosis of Alzheimer's disease dementia. The autopsy diagnoses included: Alzheimer's disease neuropathologic change, high level (A3,B3, C2) (Hyman et al., 2012); CERAD neuritic plaque score B (Mirra et al., 1991); neurofibrillary Braak stage VI/VI (Braak et al., 2006); subacute infarcts frontal, temporal, and basal ganglia; old infarct of pons; with clinicalpathological comment: "Mixed dementia, AD, and vascular. The AD component appears to predominate." The fixed brain tissue was divided into six coronal blocks of the temporal lobe that contain the entorhinal cortex, the hippocampus, and the amygdala. The orientations of the blocks correspond as closely as possible to the coordinate system of the Mai Atlas. Each block of brain tissue was scanned with a high-field 11.7T MRI scanner.

The nuclear magnetic resonance (NMR) sequence was based on a 3D multiple echo sequence (Mori and Van Zijl, 1998; Xue et al., 2001), with four echoes acquired for each excitation. The diffusion-weighted images were acquired with a field of view of typically $40 \times 30 \times 16$ mm and an imaging matrix of $160 \times 120 \times 64$, which was zero-filled to $320 \times 240 \times 128$ after the spectral data were apodized by a 10% trapezoidal function. The pixel size was native 250-micron isotropic. Eight diffusion-weighted images were acquired with different diffusion gradient directions, with *b*-values in the 1,200–1,700 s/*mm*² range. For diffusion-weighted images, a repetition time of 0.9 s, an echo time of 37 ms, and two signal averages were used, for a total imaging time of 24 h.

The MRI scanning procedure resulted in several distinct images that must themselves be aligned after imaging. For this, we developed an interactive tool for visualizing and transforming each imaged block to match with the others and for rigidly positioning the aligned blocks in the Mai Atlas coordinate system. The blocks were aligned by manually adjusting rotations and translations by small amounts until they could not be improved. The alignment was calculated by a biomedical engineer with feedback provided by a neuroanatomist. An example of our images and their alignment is shown on the left in Figure 1. The aligned blocks were labeled by a neuroanatomist for relevant medial temporal lobe structures [entorhinal cortex, subiculum, Cornu Ammonis (CA) fields, compartments of dentate gyrus, alveus]. These labels are shown as a surface reconstruction on the right in Figure 1. The superimposed lines correspond to the rostral-caudal pages of the Mai-Paxinos-Voss atlas (z-axis), with their 1 cm grid lines along the x-y axes of each page. Our definition of the entorhinal cortex (magenta) includes the medial bank of the collateral sulcus. This sulcal region (Krimer et al., 1997), also referred to as the trans entorhinal cortex, corresponds to the earliest location of AD pathology accumulation visible at autopsy (Braak and Braak, 1991). Atrophy in this region has been detected at the population level in subjects with mild cognitive impairment (Albert et al., 2011) before other changes are visible (Tward et al., 2017).

After the tissue blocks underwent high-field imaging, they were sectioned for histological examination. The tissue was sectioned at 200 μ m intervals: 20 slices of 10 μ m thickness and 5 slices of 40 μ m thickness. The thin-sliced sections are prepared with stains focused on AD pathology: Nissl, silver (Hirano method), Luxol fast blue (LFB) for myelin, and immunostained for A β (mab 6E10) and tau (PHF1). Several examples of our tau, amyloid, and myelin stained sections are shown in the results section (Figures 3-11). For future analysis, these stains are complemented by additional stains to examine the tissue for related comorbidities and other neurodegenerative disorders, such as Lewy body disease and frontotemporal dementia, including immunostains for α -synuclein, ubiquitin, TDP-43, GFAP (astrocytes), Iba1 and CD68 (microglia), collagen IV (blood vessels), and reelin (entorhinal cortex layer II neuronal protein) 13. The thick-sliced sections were used for quantitative cell and neuron counts and density studies of dendritic and synaptic markers.



3.2. Image Mapping Experiments

Below, we present several 2D experiments to demonstrate the applicability and verify the validity of our method. In each case, linear and deformable registration is performed simultaneously using gradient descent, and deformable registration is implemented using LDDMM (Equation 4a) or weighted LDDMM in an EM algorithm (Equation 7a). In our first experiment, we provide an illustrative example using a pair of simulated images that contain contrast variation, missing tissue, and artifacts.

In our second experiment, we use example histology images for a qualitative evaluation of the different components of our algorithm: contrast changes only, missing tissue only, and both together. These are approximately 1-micron resolution, with typical images about $20,000 \times 10,000$ pixels. For registration purposes, they are downsampled by averaging over a 32×32 neighborhood.

In our third experiment, we demonstrate the accuracy of our algorithm using the benchmarking dataset described in (Borovec et al., 2018). We use all seven lung tissue slides corresponding to 140 registrations between pairs of stains. While this dataset does not correspond exactly to our desired application in neuroimaging, it provides a great opportunity to calculate accuracy through landmark target registration error, which was defined in the above paper as mean landmark distance error divided by the diagonal of the image, and to compare to stateof-the-art methods. In particular, this dataset allows us to study the contrast variation component of our algorithm but not the missing data component. Using the software package described in the paper⁴, we were able to compare our registration results to alternative state-of-the-art methods, using configuration settings chosen by their authors specifically for this dataset. We sought to compare to ANTs because it scored first or second place in terms of robustness and accuracy and has been used in several neuropathology applications, including Adler et al. (2014) and

Layer	Input shape	Parameters
5×5 convolution 1	56 × 56 × 3	$5 \times 5 \times 3 \times 16 + 16 = 1,216$
ReLu 1	$56 \times 56 \times 16$	0
Max Pool 1	$56 \times 56 \times 16$	0
5×5 convolution 2	$28 \times 28 \times 16$	$5 \times 5 \times 16 \times 32 + 32 = 12,832$
ReLu 2	$28 \times 28 \times 32$	0
Max Pool 2	$28 \times 28 \times 32$	0
5×5 convolution 3	$14 \times 14 \times 64$	$5 \times 5 \times 32 \times 64 + 64 = 51,264$
ReLu 3	$14 \times 14 \times 64$	0
Max Pool 3	$14 \times 14 \times 64$	0
Flatten	$7 \times 7 \times 64$	0
Fully connected	3, 136	3, 136 × 1, 024 + 1, 024 = 3, 212, 288
ReLu	1,024	0
Dropout	1,024	Probability 0.5
Fully connected	1,024	$1,024 \times 3 + 3 = 3,075$
Cross entropy loss	3	0

bUnwarpJ because the majority of our colleagues working in microscopy use this software. We also compared to the first and second place winners of the related ANHIR challenge presented at ISBI 2019⁵, Histokat and Historeg (respectively). To give further context, we also compare to results with no registration and results with the best possible affine registration (which uses information not available to the other methods).

In our fourth experiment, we demonstrate the accuracy of our algorithm with both contrast variation and missing data and compare to the partial data local affine method of Periaswamy and Farid (2006), which has 2D source code available⁶. Using a DTI atlas set from MRICloud⁷, we performed a mapping

⁴https://github.com/Borda/BIRL

⁵https://anhir.grand-challenge.org/

⁶https://farid.berkeley.edu/downloads/code/qr_0.2.tar.gz

⁷http://www.mricloud.org



transform \mathcal{F}_{θ} , and (e) final posterior probability of atlas labels at algorithm convergence. (f,g) Show similar results for a fixed binary mask, and (h,i) show them for no mask. The identity line for contrast transforms is shown in gray. Because (e) does not render well in grayscale, we have presented it as three grayscale images in our Supplementary Material.

experiment from b0 images to trace images, between each pair of subjects in the atlas set, for a total of 56 maps. We chose the axial slice where lateral ventricles appeared the largest and calculated mean and standard deviation of Dice overlap of gray matter, white matter, and lateral ventricle. We registered whole brains to whole brains as a baseline, and, to examine missing data, we registered whole brains to hemispheres.

Our fifth and final example is to demonstrate an important application in a 3D histology pipeline, quantifying the 3D distribution of tau tangles in Alzheimer's disease. Here, we work with a sequence of transformations between 3D postmortem MRI and 2D histology slices. The sequence consists of 3D shape change (deformation and rigid positioning), adjusting for scaling (slice spacing and pixel size), position on the microscope slide (2D rigid motion), and variation in contrast (cubic intensity transformation). Rigid, scale, and deformation parameters are all jointly optimized using gradient descent. Since these transformations are combined by composition, gradients can be calculated using straightforward backpropagation (chain rule). Deformable registration uses a relatively small gradient descent step size, allowing linear transformations to be close to optimal at all times.

This pipeline quantifies the distribution of tau tangles in each 2D slice using a convolutional neural network in TensorFlow (Tward et al., 2018). Input data is 56×56 regions of interest, and the network architecture is summarized in **Table 1**. The center pixel of each region is classified as belonging to one of three classes: "tau tangles," "other tissue," or "background." A total of 2,391 training examples were identified manually by randomly choosing regions of interest and mouse-clicking in the center of tau tangles or on other tissue or background, with 8.3% positive, and neural network weights were trained using the Adam optimizer (Kingma and Ba, 2014). Every pixel in our histology data was classified using a sliding window approach as in Ciresan et al. (2012). After mapping this data into the coordinates of the Mai atlas, we report the total area of tau

tangles within the entorhinal cortex, subiculum, and CA1-3 for each atlas page. To place these numbers in context, we also report the total area of these anatomical regions (which may be affected by missing tissue) and the fraction of this area covered by tau tangles.

4. RESULTS

In this section, we present results and comments for each of our five image-registration experiments. In section 4.3, we particularly highlight the difference between our asymmetric generative statistical model and popular symmetric alternatives.

4.1. Mapping Simulated Images With Artifact and Missing Tissue

To demonstrate the method, we start with simulated images. **Figure 2** shows the atlas (**Figure 2a**) and target (**Figure 2b**). Contrast is chosen so that the atlas appears like a T1 MR brain image (darker gray matter and brighter white matter) and the target appears like a T2 MR brain image. The target also contains a bright streak artifact and missing tissue. Specifically, the atlas background has intensity 0, gray matter 1, and white matter 1.25. The target background has intensity 0, gray matter 0.9, white matter 0.675, and artifact 5. Both images have additive white Gaussian noise with standard deviation 0.05 and are blurred with a Gaussian kernel with a standard deviation of 2/3 pixels over a 5×5 pixel window.

Figures 2c,d shows that a cubic intensity transformation is sufficient to permute the order of gray and white matter, allowing for accurate matching of the cortical boundary. **Figure 2e** shows the posterior probabilities of the three atlases, shown as components of an RGB image. The pixels are correctly classified, with the atlas image in red, artifact in green, and background in blue. Note that the dark region is magenta because "atlas image" and "background" both describe the image intensity equally well.



Figure 2 also shows the failed results that occur when using a linear contrast transform only (i.e., normalized crosscorrelation), and a fixed mask (Figures 2f,g) or no mask (Figures 2h,i). Several alternative methods that use these approaches are described in section 1. With a fixed mask, the artifact is still handled appropriately. An inversion of contrast is estimated, which is appropriate within the masked region. However, missing tissue is not distinguished from normal background, so the informative "cortex/background" boundary is treated equivalently to the uninformative "cut tissue" boundary, resulting in very poor alignment. With no mask, huge distortions in shape occur as the atlas is squeezed to match the shape of the target with missing tissue and stretched to follow the bright artifact.

The number of constant-valued atlases used in addition to our template image is of importance. Because many neuroimages are approximately piecewise constant (e.g., a single intensity value for each of gray matter, white matter, and cerebrospinal fluid), using many constant atlases has the potential for overfitting. In the simulated example here, using four such atlases can describe the target image exactly by assigning a probability of zero to our non-constant atlas. This would give an undesirable solution that avoids registration entirely. In all the work shown here, we use one or two constant-valued atlases corresponding to "background" or "background and artifact." This small number of atlases leads to good quality registration without overfitting.

4.2. Mapping Histology With Missing Tissue and Different Stains

In **Figure 3**, we show results mapping a tau-stained section of the medial temporal lobe to an immediately adjacent section stained with LFB. We perform intensity transformation using a non-monotonic cubic polynomial, allowing for a swapping of brightness from gray matter (1) \rightarrow white matter (2) \rightarrow background (3) in tau to (2) \rightarrow (1) \rightarrow (3) in LFB. As a map from $\mathbb{R}^3 \rightarrow \mathbb{R}^3$, this corresponds to 60 unknown parameters (one constant, three linear, six quadratic, and 10 cubic, for each of three dimensions). This example illustrates the intensity transformation component of our algorithm in isolation.

In **Figure 4**, we map a tau-stained slice of medial temporal lobe to its neighbor, which has significant missing tissue due to damage. This illustrates the missing data component of our algorithm in isolation. In this experiment, two atlas labels are used: "tau-stained image" and "background." The



probability 0.5 in green, and low probability in blue.

top right panel shows the posterior probability that each pixel corresponds to the first label at the initialization of our algorithm, and the middle right panel shows the same probability after algorithm convergence.

In **Figure 5**, we show results mapping a tau-stained section of the medial temporal lobe to an adjacent damaged slice stained with LFB. This illustrates the intensity mapping and missing data components of our algorithm simultaneously.

4.3. Quantitative Analysis on Standardized Benchmark Datasets

Our statistical model of image formation removes the symmetry assumptions between atlas and target used in other methods (Christensen and Johnson, 2001; Avants et al., 2006, 2008), since the correspondence of histology images with 3D atlases is typically a non-invertible injection. The penalized Bayes estimator therefore computes deformations and variations only in the template image, which are compared to the histology targets via log-likelihood. For example, Equation (7a) only includes a gradient of the atlas image, whereas symmetric methods would require gradients of atlas and target. Because of missing tissue and sparse sampling, there is no 3D gradient in the target histology coordinates.

This suggests that if the template and target are of the same dimension and have similarly low noise, then the symmetric constraint is an excellent assumption and will perform well. Shown in Figure 6 are comparisons of our method to alternatives including the symmetric approach ANTs, showing the landmarkbased accuracy between pairs of histology stains in seven datasets. The results show that our method gives performance similar to that of other state-of-the-art techniques on this well-controlled dataset without missing tissue or artifacts. We obtained mean and standard deviation results for ANTs by running it locally using provided configuration files, and downloaded others from the ANHIR challenge website from which mean but not standard deviation was available. The computation time for registering these benchmark images was on average 6.6 ± 2.1 minutes when run using a standard Python implementation (FFTs computed using NumPy and interpolation computed using scipy⁸) with no explicit GPU or multi-threading support. Our system uses two 6 core Intel Xeon processors at 3.6 GHz.

⁸www.scipy.org



FIGURE 5 | Mapping across modality from tau to myelin with missing tissue. A tau-stained section through the hippocampus is mapped to a neighboring LFB-stained damaged section. The weights show the posterior probability that a given pixel has an atlas label that corresponds to our tau-stained image, not background. High probability is shown in yellow, probability 0.5 in green, and low probability in blue.



FIGURE 6 | Comparison on standard datasets. Landmark target registration error on benchmarking dataset (mean, and standard deviation if available, across all datasets in each category). To give a sense of scale, we also include error without registration and with the best possible affine registration (which uses landmark location information not available to the other methods).

It is interesting to compare the symmetric and non-symmetric approaches when we move toward scenarios where the target has significant missing or damaged tissue. Our example in Figure 5 is one such case, and in Figure 7, we compare our results to the symmetric method. The figure shows overlaps between the mapped histology section and a target section that has missing tissue. The images were generated by applying binary masks and summing the two images in RGB space. Black and white colors mean background and foreground are matching, respectively. Magenta implies that deformed atlas tissue is present but not target, and green the opposite. The symmetric method results in a roughly equal amount of magenta and green, whereas our asymmetric method results in almost entirely magenta. In situations like this one, where it is true that tissue is missing, we expect our asymmetric method to provide the desired result.

4.4. Dice Overlap for Whole Brains and Hemispheres

We show an example pair of images from our Dice overlap experiment in Figure 8 left. This illustrates the contrast



FIGURE 7 | Comparison of symmetric and asymmetric methods on a histology section with missing tissue. Deformed atlas is shown in magenta, and target image is shown in green. The two are combined by summing in RGB space. Black and white colors mean background and foreground are matching, respectively. Magenta implies deformed atlas tissue is present but not target, and green the opposite. While the symmetric method shows magenta and green, our asymmetric method shows almost entirely magenta.



differences between b0 and trace images and also illustrates our generated hemispheres. The mean and standard deviation of Dice scores across 56 mappings is shown on the right in **Figure 8** for gray matter, white matter, and lateral ventricle. Both methods examined (light vs. dark colors) perform similarly for whole-brain data. Our method shows consistently higher Dice overlaps when registering whole brains to hemispheres. This difference may be explained by our method accommodating more than one type of anomalous data or explicitly using diffeomorphic transformations.

4.5. Mapping Histology Data to Mai Atlas Coordinates

Alignments between 3D post-mortem MRI and each of our three 2D stains are shown in **Figure 9**. For each stain, we show registered MRI, intensity transformed registered MRI, and histology images. We also show the posterior probability of each atlas label ("template image," "background," "artifact") at the start and finish of the algorithm. Comparing the start to the finish demonstrates the quality of the alignment. Several regions where changes are not diffeomorphic, such as opening and closing of the ventricle, are labeled as artifact or missing tissue. The LFB stain, in particular, shows significant variation in contrast profiles from slice to slice, which is handled effectively by our method. All intensity transformations use a cubic polynomial for each of the red, green, and blue channels on each slice, which corresponds to 12 parameters.

Example results of our tau-detection algorithm, using the convolutional neural network with the architecture specified in **Table 1**, are shown in **Figure 10**. We achieve a classification accuracy of 0.996 on a test set of 500 examples that were randomly left out from the 2,391 manually classified images in our training set. The detection results are shown in **Figure 10** by overlaying the image with a red color at three different scales, with individual detections at the micron level and densities at the millimeter level.

Figure 11 shows several sections of the Mai atlas along the rostral to caudal axis in millimeters. In the same



FIGURE 9 | Mapping from MRI to three histology slices. Alignment between our 3D MRI and histology slices at 11 locations with three stains is shown. From top to bottom, boxes show data for amyloid, myelin, and tau stains. Within each box, from top to bottom: registered MRI, registered MRI intensity mapped to histology, histology, posterior probability of atlas labels at each pixel, posterior probability of atlas labels at each pixel at start of the algorithm. As in Figure 2, we use red: atlas image, blue: background, green: artifact.



FIGURE 10 | Tau-detection results. Tau-stained histology sections (top) with the result of our tangle detection (bottom, red) at three different scales, zooming out from single tangles (left) to the entire medial temporal lobe (right).



caudal-rostral axis of histology sections in Mai-Paxinos coordinates.

coordinate system, we show our post-mortem MRI with manual segmentations superimposed and our histology stains and estimated tangle density. For visualization of this sparse data in 3D, interpolation was applied between slices. To sample at a fraction *p* between slices *I* and *J*, a symmetric LDDMM transformation was computed (Avants et al., 2008), and a weighted average of images was computed from the flow (Lee et al., 2019): $(1 - p)I \circ \varphi_t^{-1} + pJ \circ \varphi_{1t}^{-1}$. Finally, **Figure 12** shows our estimated area covered by tau

Finally, **Figure 12** shows our estimated area covered by tau tangles on each page of the Mai atlas for several structures (entorhinal cortex, subiculum, and CA fields). We observe a trend of decreasing tangle concentration in the rostral to caudal

direction. Note that a 3D rendering of the segmented structures used for **Figures 11**, **12** is shown in **Figure 1**.

5. CONCLUSION

In this work, we proposed a new image-mapping method that accommodates contrast differences, missing data, and artifacts. This was achieved by formulating the imaging process as (i) an unknown shape through the action of the diffeomorphism group, (ii) an unknown change in contrast through the action of polynomial maps, and (iii) the addition of Gaussian noise. Here, (i) describes the object being imaged, and (ii) and (iii)



describe the imaging process, reflecting the distinction made by Shannon between source and channel. This model allows multimodality image-matching to be formulated as a penalized likelihood problem rather than simply the maximization of an image-similarity function. This statistical model leads naturally to the formulation of an expectation-maximization algorithm that handles missing data or artifacts. We applied this algorithm to simulated images, illustrating its effectiveness for accurate mapping and classification of image pixels and its superiority over typical alternatives. For 2D histology, we qualitatively demonstrated the effectiveness of each of our two contributions, contrast mapping and tissue classification, in isolation and simultaneously. We quantitatively demonstrated that our method gives an accuracy comparable to other state-of-the-art alternatives using a benchmarking dataset that evaluates contrast changes but not missing data. Finally, we applied this technique to the challenge of reconstructing 3D volumes from histology by mapping to post-mortem MRI. In conjunction with convolutional neural networks, this allowed us to map out the 3D distribution of tau tangles in the medial temporal lobe.

Here, we demonstrated that for $\mathbb{R} \to \mathbb{R}$ affine contrast transformations, our formulation is equivalent to normalized cross-correlation. Another popular image-similarity term is mutual information, which is invariant to all invertible transformations. Our approach can accommodate these invariances if we allow for arbitrary nonparametric transformations, which can be thought of as high-degree polynomials or as linear combinations of narrow kernel functions. In this limit, a standard result of statistical prediction results in an intensity transformation given by conditional expectation, $F(i) = E_{J|I=i}[J]$. This transformation results in a cost function with the same set of invariances as mutual information.

Typically, image registration has involved the balance between a regularization term and a data attachment term in optimization, which is characterized by a single parameter chosen to reflect the researcher's priorities. A limitation of our algorithm is that it requires more parameters: a variance for shape change (regularization) and variances of image noise, background noise, and artifact noise. However, rather than being chosen arbitrarily, these must be chosen carefully to reflect the physical characteristics of a random imaging model. The choice of polynomials here to describe intensity changes provides an efficient method for basis representation of the contrast or image variation. Further, as is typical of Expectation-Maximization algorithms, optimization in our setting can be slow and sensitive to initialization. We attempt to overcome this initialization issue using standard approaches in the registration community, including working from low resolution to high resolution and beginning with linear transformations and working toward deformations.

The scalability of the algorithm described is technically $O(N \log(N))$, where N is the number of voxels in the atlas image, owing to the Fast Fourier Transform (FFT) used for applying and inverting differential operators. In practice, however, the majority of computation time is spent in linear interpolation, which is O(N). Both linear interpolation and FFTs can be very efficiently parallelized on multicore systems or graphics-processing units.

While we compared our algorithm to two state-of-the-art methods for registration with contrast differences on a standard benchmarking dataset, this was not possible for missing data methods. To place our method in context, we briefly compare to alternative modeling approaches. Quicksilver (Yang et al., 2017), a deep learning-based method, has demonstrated robustness against anomalies that are not present in the training set. While these methods show promise, our contribution is made in the setting of explicit modeling rather than implicitly learning from large datasets. Such datasets are becoming common in whole-brain MRI but are currently lacking in histology, our intended application. Other models allow transformations to become non-smooth or noninvertible at boundaries of anomalous regions of the image (Risholm et al., 2009, 2010; Nielsen et al., 2019), which may be accurate in situations such as cut tissue. However, throughout the majority of images, transformations are diffeomorphic, and we have chosen to work in the computational anatomy random orbit model to preserve properties that are useful for morphometry in addition to registration, such as the embedding of human anatomy into a metric space. Metamorphosis-based models (Miller et al., 2002; Li et al., 2011; Nithiananthan et al., 2012) allow image intensity to vary in certain regions to match anomalies, while maskbased models (Periaswamy and Farid, 2006; Sdika and Pelletier, 2009; Vidal et al., 2009; Chitphakdithai and Duncan, 2010b) or⁹ manually or automatically ignore these anomalies. Our method leverages the strengths of these last two, modeling both non-monotonic image intensity variation and masking in a generative statistical framework.

Our contribution to AD understanding stems from the need to bridge the gap between 3D imaging such as MRI, which can be obtained in living subjects over time, and 2D histopathology, which is the technique used to make the diagnosis postmortem. While some authors have successfully registered histology to MRI in well-controlled conditions (Adler et al., 2014), we believe that the generative model proposed here, which accommodates variable contrast and missing data, will be a valuable approach for handling typical data moving forward. This earlier work, which we compare to in our benchmarking results, is well appreciated and pioneers the space via the introduction of mutual information with the symmetric methods. However, because of the importance of the noise model in our setting for digital pathology, in which the target images contain many deletions of tissue and distortions within the image plane associated with extremely sparse collections, we have not focused on the symmetry of the orbit model described in (Christensen and Johnson, 2001; Avants et al., 2006, 2008). The statistical model for image formation of Equation (5) removes symmetry assumptions between template and target, such as are associated with other methods. This is because the correspondence of histology images to the atlas is an injection. The penalized Bayes estimator therefore computes deformations and variations only in the atlas image, as shown by the necessary maximizer condition of Equation (7a). This paper emphasizes the statistical estimation model in which the atlas is an idealized representation, i.e., the atlas is densely sampled to 100 microns isotropically, while the target has nonuniform sparse sampling to 1-2 mm with significant noise associated to nuisance dimensions that are unique to the histological preparation. An additional advantage of our statistical likelihood interpretation for estimating missing or distorted image data is that it allows us to perform Bayesian calculations such as are used in multi-atlas interpretation (Tang et al., 2012; Wang et al., 2012) and Bayesian segmentation (Tang et al., 2013).

This work advances the field of brain mapping in two important ways. First, it moves to exploit statistical models of image formation under the key assumption that the measured targets are sparse and noisy samples of the dense atlases. No symmetric property is available because mappings from histology slices to 3D atlases are injective: while there may be only one point in the atlas corresponding to each point in the target, the atlas is far more densely sampled. Second, this method accommodates mapping between images taking values in arbitrary dimensions in the presence of missing tissue and artifacts. This allows accurate brain mapping to expand from well-controlled clinical imaging to the massive diversity of neuroscience data. For example, in the mouse community, accurate image-mapping between Nissl-stained tissue and microscopy with multiple fluorophores is commonly required in the presence of variably dissected or damaged tissue. We are currently applying these techniques to CLARITY (Chung et al., 2013; Epp et al., 2015) and iDISCO (Renier et al., 2014) images in mouse and rat (Branch et al., 2018), serially sectioned mouse as part of the BRAIN Initiative Cell Census Network (Lee et al., 2018) and revisiting older datasets where images were excluded due to artifacts or damaged tissue.

DATA AVAILABILITY STATEMENT

The datasets analyzed and generated for this study can be found on the Center for Imaging Science's FTP server¹⁰. This repository contains four datasets and a README file that describes which datasets were used to make which figures. In particular, we include images from our phantom experiment (**Figure 2**), *ex-vivo* MRI images and their manual segmentations (**Figures 1**, **9**, **11**), low-resolution histology images that were used for registration (**Figures 3**, **4**, **5**, **7**, **9**, **11**), and high-resolution histology images with detected tau tangle probability (**Figures 10–12**). Images from the benchmarking dataset can be acquired at¹¹.

AUTHOR CONTRIBUTIONS

DT and MM designed the algorithm. DT wrote the resulting code and performed the experiments. TB performed the manual annotations. YK, JP, and JT prepared the histology images. ZH, JP, and SM prepared the post-mortem MRI. DT, SM, MA, JT, and MM conceived of this study and contributed to the writing and revising of the manuscript.

FUNDING

This work was supported by the National Institutes of Health (NIH) (www.nih.gov) grants P41EB015909 (SM), R01NS086888 (SM), R01EB020062 (MM), R01NS102670 (MM), U19AG033655 (MM), and R01MH105660 (MM), the National Science Foundation (NSF) (www.nsf.gov) 16-569 NeuroNex contract 1707298 (MM), the Computational Anatomy Science Gateway (DT and MM) as part of the Extreme Science and Engineering Discovery Environment (XSEDE Towns et al., 2014), which is supported by the NSF grant ACI1548562, Johns Hopkins University Alzheimer's Disease Research Center with NIH grant P50AG05146 (MA), the Dana Foundation's (www. dana.org) clinical neuroscience research program (MA), the BrightFocus Foundation (www.brightfocus.org) (JT), and the Kavli Neuroscience Discovery Institute (kavlijhu.org) supported by the Kavli Foundation (www.kavlifoundation.org) (DT, MM, and JT). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins. 2020.00052/full#supplementary-material

⁹https://github.com/armaneshaghi/LesionFilling_example

¹⁰ftp://ftp.cis.jhu.edu/pub/data_share/tward_frontiers_neuroscience_2019/ ¹¹http://cmp.felk.cvut.cz/~borovji3/?page=dataset

REFERENCES

- Adler, D. H., Pluta, J., Kadivar, S., Craige, C., Gee, J. C., Avants, B. B., et al. (2014). Histology-derived volumetric annotation of the human hippocampal subfields in postmortem MRI. *Neuroimage* 84, 505–523. doi: 10.1016/j.neuroimage.2013.08.067
- Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., et al. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the national institute on aging-Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement.* 7, 270–279. doi: 10.1016/j.jalz.2011. 03.008
- Ardekani, S., Jain, A., Jain, S., Abraham, T. P., Abraham, M. R., Zimmerman, S., et al. (2012). "Matching sparse sets of cardiac image cross-sections using large deformation diffeomorphic metric mapping algorithm," in *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges*, eds O. Camara, E. Konukoglu, M. Pop, K. Rhode, M. Sermesant, and A. Young, 7085 in Lecture Notes in Computer Science (Berlin; Heidelberg:Springer), 234–243.
- Ashburner, J. (2009). Computational anatomy with the SPM software. Magn. Reson. Imaging 27, 1163–1174. doi: 10.1016/j.mri.2009. 01.006
- Ashburner, J., and Friston, K. J. (2007). "Computational anatomy," in *Statistical Parametric Mapping The Analysis of Functional Brain Images*, eds K. J. Friston, J. Ashburner, S. J. Kiebel, T. E. Nichols, and W. D. Penny (London; Burlington, MA; San Diego, CA: Academic Press), 49–100.
- Avants, B. B., Epstein, C. L., Grossman, M., and Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med. Image Anal.* 12, 26–41. doi: 10.1016/j.media.2007.06.004
- Avants, B. B., Grossman, M., and Gee, J. C. (2006). "Symmetric diffeomorphic image registration: evaluating automated labeling of elderly and neurodegenerative cortex and frontal lobe," in *Proceedings of the Third International Conference on Biomedical Image Registration, WBIR'06* (Berlin; Heidelberg: Springer-Verlag), 50–57.
- Avants, B. B., Tustison, N. J., Stauffer, M., Song, G., Wu, B., and Gee, J. C. (2014). The insight toolkit image registration framework. *Front. Neuroinformatics* 8:44. doi: 10.3389/fninf.2014.00044
- Bashiri, F., Baghaie, A., Rostami, R., Yu, Z., and D'Souza, R. (2019). Multi-modal medical image registration with full or partial data: a manifold learning approach. J. Imaging 5:5. doi: 10.3390/jimaging50 10005
- Beg, M. F., Miller, M. I., Trouvé, A., and Younes, L. (2005). Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *Int. J. Comput. Vis.* 61, 139–157. doi: 10.1023/B:VISI.0000043755.93 987.aa
- Borovec, J., Munoz-Barrutia, A., and Kybic, J. (2018). "Benchmarking of image registration methods for differently stained histological slides," in 2018 25th IEEE International Conference on Image Processing (ICIP) (Athens: IEEE), 3368–3372.
- Braak, H., Alafuzoff, I., Arzberger, T., Kretzschmar, H., and Del Tredici, K. (2006). Staging of alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. *Acta Neuropathol.* 112, 389–404. doi: 10.1007/s00401-006-0127-z
- Braak, H., and Braak, E. (1991). Neuropathological stageing of Alzheimerrelated changes. Acta Neuropathol. 82, 239–259. doi: 10.1007/bf00 308809
- Branch, A. E., Tward, D. J., Chandrashekar, V., Miller, M., Vogelstein, J. T., and Gallagher, M. (2018). "Registration methodology for cleared rodent brain tissue," in *Society for Neuroscience Annual Meeting 2018* (San Diego, CA), 611.12/ LLL39.
- Brett, M., Leff, A. P., Rorden, C., and Ashburner, J. (2001). Spatial normalization of brain images with focal lesions using cost function masking. *Neuroimage* 14, 486–500. doi: 10.1006/nimg.2001.0845
- Chitphakdithai, N., and Duncan, J. S. (2010a). "Non-rigid registration with missing correspondences in preoperative and postresection brain images," in

International Conference on Medical Image Computing and Computer-Assisted

- Intervention (Beijing: Springer), 367–374. Chitphakdithai, N., and Duncan, J. S. (2010b). "Pairwise registration of images with missing correspondences due to resection," in 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro (Rotterdam: IEEE), 1025–1028.
- Christensen, G. E., and Johnson, H. J. (2001). Consistent image registration. *IEEE Trans. Med. Imaging* 20, 568–582. doi: 10.1109/42. 932742
- Chuang, N., Mori, S., Yamamoto, A., Jiang, H., Ye, X., Xu, X., et al. (2011). An MRI-based atlas and database of the developing mouse brain. *Neuroimage* 54, 80–89. doi: 10.1016/j.neuroimage.2010.07.043
- Chung, K., Wallace, J., Kim, S.-Y., Kalyanasundaram, S., Andalman, A. S., Davidson, T. J., et al. (2013). Structural and molecular interrogation of intact biological systems. *Nature* 497:332. doi: 10.1038/nature 12107
- Ciresan, D., Giusti, A., Gambardella, L. M., and Schmidhuber, J. (2012). "Deep neural networks segment neuronal membranes in electron microscopy images," in *Advances in Neural Information Processing Systems* (Lake Tahoe, NV), 2843–2851.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. J. R. Stat. Soc. Ser. B 39, 1–38. doi: 10.1111/j.2517-6161.1977.tb01600.x
- Durrleman, S., Pennec, X., Trouvé, A., and Ayache, N. (2008). "A forward model to build unbiased atlases from curves and surfaces," in *Proceedings of the International Workshop on the Mathematical Foundations of Computational Anatomy MFCA2008* (New York, NY).
- Epp, J. R., Niibori, Y., Hsiang, H.-L., Mercaldo, V., Deisseroth, K., Josselyn, S. A., et al. (2015). Optimization of clarity for clearing whole brain and other intact organs. *eneuro* 2, 1–15. doi: 10.1523/ENEURO.0022-15.2015
- Evans, A. C., Collins, D. L., Mills, S., Brown, E., Kelly, R., and Peters, T. M. (1993). "3D statistical neuroanatomical models from 305 MRI volumes," in *Nuclear Science Symposium and Medical Imaging Conference*, 1993., 1993 IEEE Conference Record. (San Francisco, CA: IEEE), 1813–1817.
- Grenander, M., and Miller, M. (2007). Pattern Theory: From Representation to Inference. Oxford: Oxford Press.
- Grenander, U. (1994). General Pattern Theory. Oxford: Oxford Science Publications.
- Grenander, U., and Miller, M. I. (1998). Computational anatomy: an emerging discipline. *Q. Appl. Math.* 56, 617–694. doi: 10.1090/qam/16 68732
- Heinrich, M. P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F. V., Brady, S. M., et al. (2012). Mind: modality independent neighbourhood descriptor for multi-modal deformable registration. *Med. Image Anal.* 16, 1423–1435. doi: 10.1016/j.media.2012.05.008
- Hyman, B. T., Phelps, C. H., Beach, T. G., Bigio, E. H., Cairns, N. J., Carrillo, M. C., et al. (2012). National Institute on aging–Alzheimer's association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimer's Dement.* 8, 1–13. doi: 10.1016/j.jalz.2011.10.007
- Kingma, D. P., and Ba, J. (2014). Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980.
- Krimer, L. S., Hyde, T. M., Herman, M. M., and Saunders, R. C. (1997). The entorhinal cortex: an examination of cyto-and myeloarchitectonic organization in humans. *Cereb. Cortex* 7, 722–731. doi: 10.1093/cercor/ 7.8.722
- Kulason, S., Tward, D. J., Brown, T., Sicat, C. S., Liu, C.-F., Ratnanather, J. T., et al. (2019). Cortical thickness atrophy in the transentorhinal cortex in mild cognitive impairment. *NeuroImage* 21:101617. doi: 10.1016/j.nicl.2018. 101617
- Kutten, K. S., Charon, N., Miller, M., Ratnanather, J., Matelsky, J., Baden, A. D., et al. (2017). A large deformation diffeomorphic approach to registration of clarity images via mutual information. arXiv:1612.00356 275–282.
- Lee, B. C., Tward, D. J., Mitra, P. P., and Miller, M. I. (2018). On variational solutions for whole brain serial-section histology using the computational anatomy random orbit model. arXiv preprint arXiv:1802.03453. doi: 10.1371/journal.pcbi.1006610

- Lee, B. C., Tward, D. J., Wei, J., Tipre, D., Weiss, R. G., Miller, M. I., et al. (2019). "Diffeomorphic upsampling of serially acquired sparse 2D crosssections in cardiac MRI," in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (Berlin), 4491–4495. doi: 10.1109/EMBC.2019.8856317
- Li, X., Long, X., Laurienti, P., and Wyatt, C. (2011). Registration of images with varying topology using embedded maps. *IEEE Trans. Med. Imaging* 31, 749–765. doi: 10.1109/TMI.2011.2178609
- Mai, J., and Paxinos, G. (2011). *The Human Nervous System*. San Diego, CA; London: Academic Press.
- Mattes, D., Haynor, D. R., Vesselle, H., Lewellen, T. K., and Eubank, W. (2003). PET-CT image registration in the chest using free-form deformations. *IEEE Trans. Med. Imaging* 22, 120–128. doi: 10.1109/TMI.2003.809072
- Miller, M. I., Arguillère, S., Tward, D. J., and Younes, L. (2018). Computational anatomy and diffeomorphometry: a dynamical systems model of neuroanatomy in the soft condensed matter continuum. *Wiley Interdisc. Rev.* 10:e1425. doi: 10.1002/wsbm.1425
- Miller, M. I., Christensen, G. E., Amit, Y., and Grenander, U. (1993). Mathematical textbook of deformable neuroanatomies. *Proc. Natl. Acad. Sci. U.S.A.* 90, 11944–11948. doi: 10.1073/pnas.90.24.11944
- Miller, M. I., Faria, A. V., Oishi, K., and Mori, S. (2013a). Highthroughput neuro-imaging informatics. *Front. Neuroinformatics* 7:31. doi: 10.3389/fninf.2013.00031
- Miller, M. I., Ratnanather, J. T., Tward, D. J., Brown, T., Lee, D. S., Ketcha, M., et al. (2015). Network neurodegeneration in Alzheimer's disease via MRI based shape diffeomorphometry and high-field atlasing. *Front. Bioeng. Biotechnol.* 3:54. doi: 10.3389/fbioe.2015.00054
- Miller, M. I., Trouve, A., and Younes, L. (2002). On the metrics and Euler-lagrange equations of computational anatomy. *Annu. Rev. Biomed. Eng.* 4, 375–405. doi: 10.1146/annurev.bioeng.4.092101.125733
- Miller, M. I., Trouvé, A., and Younes, L. (2014). Diffeomorphometry and geodesic positioning systems for human anatomy. *Technology* 1:36. doi: 10.1142/S2339547814500010
- Miller, M. I., Younes, L., Ratnanather, J. T., Brown, T., Trinh, H., Postell, E., et al. (2013b). The diffeomorphometry of temporal lobe structures in preclinical alzheimer's disease. *NeuroImage* 3, 352–360. doi: 10.1016/j.nicl.2013.09.001
- Mirra, S. S., Gearing, M., and Nash, F. (1997). Neuropathologic assessment of Alzheimer's disease. *Neurology* 49(3 Suppl. 3), S14–S16. doi: 10.1212/WNL49.3_Suppl_3.S14
- Mirra, S. S., Heyman, A., McKeel, D., Sumi, S. M., Crain, B. J., Brownlee, L. M., et al. (1991). The consortium to establish a registry for Alzheimer's disease (cerad) part II. standardization of the neuropathologic assessment of Alzheimer's disease. *Neurology* 41:479. doi: 10.1212/wnl.41.4.479
- Mori, S., Oishi, K., Faria, A., and Zijl, P. (2005). *MRI Atlas of Human White Matter*. London; Burlington, MA; San Diego, CA: Elsevier.
- Mori, S., Oishi, K., Faria, A. V., and Miller, M. I. (2013). Atlas-based neuroinformatics via MRI: harnessing information from past clinical cases and quantitative image analysis for patient care. *Annu. Rev. Biomed. Eng.* 15, 71–92. doi: 10.1146/annurev-bioeng-071812-152335
- Mori, S., and Van Zijl, P. C. (1998). A motion correction scheme by twin-echo navigation for diffusion-weighted magnetic resonance imaging with multiple RF echo acquisition. *Magn. Reson. Med.* 40, 511–516. doi: 10.1002/mrm.1910400403
- Nielsen, R. K., Darkner, S., and Feragen, A. (2019). "Topaware: topology-aware registration," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Shenzhen: Springer), 364–372.
- Nithiananthan, S., Schafer, S., Mirota, D. J., Stayman, J. W., Zbijewski, W., Reh, D. D., et al. (2012). Extra-dimensional demons: a method for incorporating missing tissue in deformable image registration. *Med. Phys.* 39, 5718–5731. doi: 10.1118/1.4747270
- Periaswamy, S., and Farid, H. (2006). Medical image registration with partial data. Med. Image Anal. 10, 452–464. doi: 10.1016/j.media.2005.03.006
- Pichat, J., Iglesias, J. E., Yousry, T., Ourselin, S., and Modat, M. (2018). A survey of methods for 3D histology reconstruction. *Med. Image Anal.* 46, 73–105. doi: 10.1016/j.media.2018.02.004
- Pluim, J. P., Maintz, J. B., and Viergever, M. A. (2003). Mutual-informationbased registration of medical images: a survey. *IEEE Trans. Med. Imaging* 22, 986–1004. doi: 10.1109/TMI.2003.815867

- Renier, N., Wu, Z., Simon, D. J., Yang, J., Ariel, P., and Tessier-Lavigne, M. (2014). iDISCO: a simple, rapid method to immunolabel large tissue samples for volume imaging. *Cell* 159, 896–910. doi: 10.1016/j.cell.2014.10.010
- Risholm, P., Samset, E., Talos, I.-F., and Wells, W. (2009). "A non-rigid registration framework that accommodates resection and retraction," in *International Conference on Information Processing in Medical Imaging* (Williamsburg, VA: Springer), 447–458.
- Risholm, P., Samset, E., and Wells III, W. (2010). "Validation of a nonrigid registration framework that accommodates tissue resection," in *Medical Imaging 2010: Image Processing, Vol. 7623.* (San Diego, CA: International Society for Optics and Photonics), 762319.
- Sdika, M., and Pelletier, D. (2009). Nonrigid registration of multiple sclerosis brain images using lesion inpainting for morphometry or lesion mapping. *Hum. Brain Mapp.* 30, 1060–1067. doi: 10.1002/hbm.20566
- Stefanescu, R., Commowick, O., Malandain, G., Bondiau, P.-Y., Ayache, N., and Pennec, X. (2004). "Non-rigid atlas to subject registration with pathologies for conformal brain radiotherapy," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Saint-Malo: Springer), 704–711.
- Sunkin, S. M., Ng, L., Lau, C., Dolbeare, T., Gilbert, T. L., Thompson, C. L., et al. (2013). Allen brain atlas: an integrated spatio-temporal portal for exploring the central nervous system. *Nucleic Acids Res.* 41, D996–D1008. doi: 10.1093/nar/gks1042
- Talairach, J., and Szikla, G. (1980). "Application of stereotactic concepts to the surgery of epilepsy," in Advances in Stereotactic and Functional Neurosurgery, Vol. 4, eds F. J. Gillingham, J. Gybels, E. Hitchcock, G. F. Rossi, and G. Szikla (Vienna: Springer), 35–54.
- Tang, X., Mori, S., and Miller, M. (2012). "Segmentation via multi-atlas LDDMM," in MICCAI 2012 Workshop on Multi-Atlas Labeling (Nice), 123–133.
- Tang, X., Oishi, K., Faria, A. V., Hillis, A. E., Albert, M. S., Mori, S., et al. (2013). Bayesian parameter estimation and segmentation in the multi-atlas random orbit model. *PLoS ONE* 8:e65591. doi: 10.1371/journal.pone.0065591
- Thompson, P. M., and Toga, A. W. (2002). A framework for computational anatomy. Comput. Visual. Sci. 5, 13–34. 65591
- Toga, A. W., and Thompson, P. M. (2001). Maps of the brain. *Anat Rec.* 265, 37–53. doi: 10.1002/ar.1057
- Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., et al. (2014). Xsede: accelerating scientific discovery. *Comput. Sci. Eng.* 16, 62–74. doi: 10.1109/MCSE.2014.80
- Tward, D. J., Brown, T., Patel, J., Kageyama, Y., Mori, S., Troncoso, J. C., et al. (2018). Quantification of 3D tangle distribution in medial temporal lobe using multimodal image registration and convolutional neural networks. *Alzheimer's Dement.* 14:P57. doi: 10.1016/j.jalz.2018.06.2126
- Tward, D. J., Sicat, C. S., Brown, T., Bakker, A., Gallagher, M., Albert, M., et al. (2017). Entorhinal and transentorhinal atrophy in mild cognitive impairment using longitudinal diffeomorphometry. *Alzheimer's Dement.* 9, 41–50. doi: 10.1016/j.dadm.2017.07.005
- Vialard, F.-x., Risser, L., Holm, D., and Rueckert, D. (2011). "Diffeomorphic atlas estimation using Kärcher mean and geodesic shooting on volumetric images Kärcher mean," in *Medical Image Understanding* and Analysis, eds W. R. Crum and G. Penney (London).
- Vidal, C., Hewitt, J., Davis, S., Younes, L., Jain, S., and Jedynak, B. (2009). "Template registration with missing parts: application to the segmentation of *M. tuberculosis* infected lungs," in 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro (Boston, MA: IEEE), 718–721.
- Wachinger, C., and Navab, N. (2012). Entropy and laplacian images: Structural representations for multi-modal registration. *Med. Image Anal.* 16, 1–17. doi: 10.1016/j.media.2011.03.001
- Wang, H., Suh, J. W., Das, S. R., Pluta, J. B., Craige, C., and Yushkevich, P. A. (2012). Multi-atlas segmentation with joint label fusion. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 611–623. doi: 10.1109/TPAMI.2012.143
- Wu, J., and Tang, X. (2018). "Fast diffeomorphic image registration via GPUbased parallel computing: an investigation of the matching cost function," in *Proceedings of SPIE Medical Imaging (SPIE-MI)* (Orlando, FL).
- Xue, R., Sawada, M., Goto, S., Hurn, P. D., Traystman, R. J., van Zijl, P. C., et al. (2001). "Rapid three-dimensional diffusion mri facilitates the study of acute stroke in mice," *Magn. Reson. Med.* 46, 183–188. doi: 10.1002/ mrm.1174

- Yang, X., Kwitt, R., Styner, M., and Niethammer, M. (2017). Quicksilver: fast predictive image registration-a deep learning approach. *NeuroImage* 158, 378– 396. doi: 10.1016/j.neuroimage.2017.07.008
- Yoo, T. S., Ackerman, M. J., Lorensen, W. E., Schroeder, W., Chalana, V., Aylward, S., et al. (2002). Engineering and algorithm design for an image processing API: a technical report on ITK-the insight toolkit. *Stud. Health Technol. Inform.* 85, 586–592.
- Younes, L., Albert, M., Miller, M. I., et al. (2014). Inferring changepoint times of medial temporal lobe morphometric change in preclinical Alzheimer's disease. *NeuroImage* 5, 178–187.
- Younes, L., Albert, M. S., Moghekar, A., Soldan, A., Pettigrew, C., and Miller, M. I. (2019). Identifying changepoints in biomarkers during the preclinical phase of AD. *Front. Aging Neurosci.* 11:74. doi: 10.3389/fnagi.2019. 00074
- Zacharaki, E. I., Shen, D., Lee, S.-K., and Davatzikos, C. (2008). Orbit: A multiresolution framework for deformable registration of brain tumor images. *IEEE Trans. Med. Imaging* 27, 1003–1017. doi: 10.1109/TMI.2008.916954
- Zhang, D. P., Risser, L., Vialard, F.-X., Edwards, P., Metz, C., Neefjes, L., et al. (2010a). Coronary motion estimation from CTA using probability atlas and diffeomorphic registration. *Med. Imaging Augment. Real.* 6326, 78–87. doi: 10.1007/978-3-642-15699-1_9

Zhang, Y., Zhang, J., Oishi, K., Faria, A. V., Jiang, H., Li, X., et al. (2010b). Atlas-guided tract reconstruction for automated and comprehensive examination of the white matter anatomy. *NeuroImage* 52, 1289–1301. doi: 10.1016/j.neuroimage.2010. 05.049

Conflict of Interest: SM and MM own Anatomy Works, with SM serving as its CEO. This arrangement is being managed by Johns Hopkins University in accordance with its conflict of interest policies.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Tward, Brown, Kageyama, Patel, Hou, Mori, Albert, Troncoso and Miller. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.