Diffeomorphic registration with intensity transformation and missing data: Application to 3D digital pathology of Alzheimer's disease

Daniel Tward[1,2*], Timothy Brown[1], Yusuke Kageyama[3,¤], Jaymin Patel[4], Zhipeng Hou[5], Susumu Mori[5], Marilyn Albert[6], Juan Troncoso[3,6], Michael Miller[1,2,4]

1 Center for Imaging Science, Johns Hopkins University, Baltimore, Maryland, USA.
2 Kavli Neuroscience Discovery Institute, Departments of Neuroscience and Biomedical Engineering, The Johns Hopkins University, Baltimore, Maryland, USA.
3 Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.
4 Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland, USA.
5 Department of Radiology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.
6 Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

¤ Current Affiliation: Molecular Neuroscience Research Center, Shiga University of Medical Science, Otsu, Shiga, Japan.
* dtward@cis.jhu.edu

Abstract

This paper examines the problem of diffeomorphic image mapping in the presence of differing image intensity profiles and missing data. Our motivation comes from the problem of aligning 3D brain MRI with 100 micron isotropic resolution, to histology sections with 1 micron in plane resolution. Multiple stains, as well as damaged, folded, or missing tissue are common in this situation. We overcome these challenges by introducing two new concepts. Cross modality image matching is achieved by jointly estimating polynomial transformations of the atlas intensity, together with pose and deformation parameters. Missing data is accommodated via a multiple atlas selection procedure where several atlases 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 show results for 3D reconstruction of digital pathology and MRI in standard atlas coordinates. In conjunction with convolutional neural networks, we quantify the 3D density distribution of tauopathy throughout the medial temporal lobe of an Alzheimer's disease postmortem specimen.

Author summary

Our work in Alzheimer's disease (AD) is attempting to connect histopathology at autopsy and longitudinal clinical magnetic resonance imaging (MRI), combining the strengths of each modality in a common coordinate system. We are bridging this gap by

using post mortem high resolution MRI to reconstruct digital pathology in 3D. This image registration problem is challenging because it combines images from different modalities in the presence of missing tissue and artifacts. We overcome this challenge by developing a new registration technique that simultaneously classifies each pixel as "good data" / "missing tissue" / "artifact", learns a contrast transformation between modalities, and computes deformation parameters. We name this technique "(D)eformable (R)egistration and (I)ntensity (T)ransformation with (M)issing (D)ata", pronounced as "Dr. It, M.D.". In conjunction with convolutional neural networks, we use this technique to map the three dimensional distribution of tau tangles in the medial temporal lobe of an AD postmortem specimen.

Introduction

High throughput neuroinformatics is emerging in neuroscience [1, 2]. 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, including Tailarach coordinates [3], Montreal Neurological Institute (MNI) [4], and Mori's diffusion tensor imaging (DTI) white matter atlases [5], which define the locations of neuroanatomical structures as well as important structural and functional properties such as volume, shape, blood oxygen-level dependent (BOLD) signals, etc. At micron and meso-scales there are several atlases including Mori's and Allen brain atlas [6, 7] with their associated region and cell-types.

High-throughput image analysis relies on some form of dense brain mapping. At the millimeter scale there have been several approaches to computational anatomy with deformable templates [8–13], atlas estimation [14, 15], and applications in white matter [16] or even cardiac imaging [17, 18] . Many of these algorithms have been extended to micron scales, such as for CLARITY [19–21]. However, most of these datasets and the methods used to analyze them are based on high quality clinical images. At the frontier of neuroscience are images that are less controlled. In this work we focus on digital pathology, where micron thick tissue slices are prone to damage and subsets which are completely missing, and where a host of stains can result in a multitude of different image intensities. To bring digital pathology to the era of high throughput neuroinformatics, brain mapping algorithms need to be expanded to handle missing data and multiple modalities with perhaps radically different contrast profiles.

The community has developed a host of techniques for addressing these situations partially. Image masking allows algorithms to ignore missing or unreliable data, and similarity functions such as normalized cross correlation [22, 23] or mutual information [24] allow mapping between different image contrasts to an extent. These tools are available in standard software packages such as the Insight Toolkit (ITK) [25]. However, modern image mapping techniques such as multi atlas [26] or Bayesian segmentation [27] require not just similarity functions, but a complete generative model that describes the likelihood of each observed image voxel.

In this work we develop a generative probabilistic model that accounts for differences in contrast and missing or censored data. We use the random orbit model of Computational Anatomy [9, 28] in which the space of histological images is an orbit of exemplar templates under smooth coordinate transformation, as well as transformation of contrast. The models for coordinate transformations are taken from diffeomorphometry [29, 30]. We wish to accommodate differing contrasts associated to different histological staining such tau, amyloid, myelin, and a variety of contrasts such as Nissl and fluorescence at the meso- and micro-scale. To achieve this, we model the space of contrasts manifest in the orbit of observed images as polynomial functions of

the template. In the case of first order polynomials, this gives affine models which we show reduce to the normalized cross-correlation cost function. Application of higher order polynomials can describe non-monotone transformations which swap the order of intensities. We name this mapping framework, "(D)eformable (R)egistration and (I)ntensity (T)ransformation", pronounced as "Dr. It".

To accommodate effects such as folded or missing tissue, we include additional homogeneous atlases, and model each pixel in an observed image as a realization of one transformed atlas from this family. Pixels corresponding to missing tissue, data deletion, or global censoring are interpreted predominantly via the "background only" atlas. Pixels corresponding to anomalous intensities are interpreted predominantly via the "artifact" atlas. The unknown atlas label at each pixel is interpreted as missing data, and its conditional mean is estimated using the Expectation Maximization algorithm [31]. The conditional mean is interpreted as the posterior probability of a particular atlas in the multi atlas random orbit model. We name the version of our algorithm that includes missing data "(D)eformable (R)egistration and (I)ntensity (T)ransformation with (M)issing (D)ata", pronounced as "Dr. It, M.D.".

In this paper, our framework is studied using simulated examples, and those from digital pathology and MRI. 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. Tangles are detected from histology images using convolutional neural networks, registered to ex vivo MRI, and mapped to the standard coordinates of the Mai Paxinos Voss atlas [32].

Methods

Ethics Statement

The Johns Hopkins Medicine Office of Human Subjects Research Institutional Review Boards determined that this work does not constitute human subjects research under the DHHS or FDA regulations (IRB00101384).

The contrast transformation problem

Our generative model for the space of observables or target images builds upon the deformable templates of [28, 33]. Observed images are transformations of collections of exemplars or atlas images $I : X \subset \mathbb{R}^3 \to \mathbb{R}^N$, which may be single valued ($N = 1$) such as a T1 MRI, or multi-valued such as red green blue (RGB, $N = 3$). The atlases experience two kinds of transformations: deformations on the background space $\varphi : X \to X$, and transformations of the image intensity $F : \mathbb{R}^N \to \mathbb{R}^M$. The mapping F depends on imaging instrumentation and tissue properties. Considering $N \neq M$ can accommodate mappings between many different modalities.

We define the deformations of coordinates as diffeomorphisms, $\varphi \in \mathit{Diff}$, generated using flows as in [9], $\varphi = \phi_1, \dot{\phi}_t = v_t \circ \phi_t, t \in [0, 1], \phi_0 = id$, v_t the Eulerian vector fields of the flow. The mapping of image intensity is defined parametrically, $F_\theta : \mathbb{R}^N \to \mathbb{R}^M$, $\theta \in \Theta$ unknown in some parameter set. In this work we consider F as polynomials with Θ the set of coefficients.

The random imaging model takes the observable images $J(x), x \in X$ as a conditionally Gaussian random field having mean given by the transformed template, and white noise variance σ^2:

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

The log-likelihood, as a function of the unknown parameters is

$$\ell(J;\varphi,\theta) = \frac{1}{2\sigma^2}\|J - F_\theta(I\circ\varphi^{-1})\|_{L_2}^2 \ . \tag{1}$$

While the dimension of θ is finite, that of the diffeomorphism is infinite. We therefore use penalized methods, introducing a Sobolev norm on the vector fields $\|v_t\|_V^2$ as a running penalty. We force $v \in V$ to be a reproducing kernel Hilbert space with kernel K defined through the differential operator $A : v \in V \mapsto Av \in V^*$, V^* the dual-space of smooth vector fields V. Then the norm written as generalized function integration becomes

$$\|v\|_V^2 = \int_X Av \cdot v dx \ .$$

The penalized likelihood becomes for $\varphi = \int_0^1 v_t \circ \phi_t dt + id$, and σ_R^2 a regularization parameter:

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

To calculate the penalized maximum likelihood estimators we assume the mapping F_θ differentiable with $\frac{d}{d\theta}F_\theta$ an $M \times B$ vector, B being the number of basis functions in our polynomial.

Theorem 1. *Minimizers of he penalized likelihood of* (2) *satisfy*

$$Av_t = \frac{1}{\sigma^2} D\left(F_\theta(I\circ\varphi_t^{-1})\right)^T \left(J(\varphi_{1t}^{-1}) - F_\theta(I\circ\varphi_t^{-1})\right)|D\varphi_{1t}^{-1}| \tag{3a}$$

$$\int_X J^T(x)\frac{d}{d\theta}F_\theta(I\circ\varphi_1^{-1}(x))dx = \int_X F_\theta(I\circ\varphi_1^{-1}(x))^T\frac{d}{d\theta}F_\theta(I\circ\varphi_1^{-1}(x))dx \tag{3b}$$

The first equation (3a) is the necessary condition for the stationary solution with respect to the deformation controlled by the vector field, the original LDDMM equation of Beg [34], where $\varphi_{1t} = \varphi_t \circ \varphi_1^{-1}$ is a mapping from time 1 to time t. The equation (3a) is computed by application of the chain rule. Define for notational convenience $\tilde{I} \doteq I \circ \varphi^{-1}$. First the derivative of E with respect to $F_\theta(\tilde{I})$, second $F_\theta(\tilde{I})$ with respect to $\tilde{I}$, and third $\tilde{I}$ with respect to the deformation field. The first and second steps are combined as

$$\frac{\partial E}{\partial \tilde{I}} = \frac{1}{\sigma^2}(F_\theta(\tilde{I}) - J)^T DF_\theta(\tilde{I}) \tag{4}$$

The third step is discussed in [34]. This equation can be solved using a standard gradient descent approach. The equation (3b) is the necessary condition with respect to the contrast or photometric parameters. Here we consider F_θ as a linear combination of B polynomial basis functions, and so (3b) is a linear system solved exactly at each iteration of gradient descent.

Minimizing over both θ and transformation parameters means the result of registration will be independent of the family of transformations indexed by θ. Alternatively, we can consider minimizing over θ first, leading to an invariant cost function

$$\text{Cost}_{\text{invariant to } F_\theta \forall\theta\in\Theta}(I) = \min_\theta \frac{1}{2\sigma^2}\|F_\theta(I) - J\|_{L_2}^2$$

Corollary (Unknown Affine Transformations Correspond to Normalized Cross Correlation Squared)**.** *When F is an affine map from $\mathbb{R} \to \mathbb{R}$, our registration results*

will be invariant to affine transformations of our atlas, and (3b) *can be solved analytically. For* $N = M = 1$ *and* $a, b \in \mathbb{R}$, $F_{a,b}(t) = at + b$, *minimizing* $\frac{1}{2\sigma^2}\|aI \circ \varphi^{-1} + b - J\|_{L_2}^2$ *is equivalent to maximizing the normalized cross correlation squared of* $I \circ \varphi^{-1}$ *with* J.

Proof. For any fixed φ, optimal values of a, b can be found via a standard linear least squares estimation result, which gives $a = \mathrm{Cov}(\tilde{I}, J)/\mathrm{Var}(\tilde{I})$ and $b = \bar{J} - a\bar{\tilde{I}}$ where $\bar{\cdot}$ corresponds to the expected value and expectation is taken by averaging over all voxels in the images.

Plugging these into our L_2 cost gives

$$\begin{aligned}
&\frac{1}{2\sigma^2}\int_X \left|\frac{\mathrm{Cov}(\tilde{I}, J)}{\mathrm{Var}(\tilde{I})}\tilde{I}(x) + \bar{J} - \frac{\mathrm{Cov}(\tilde{I}, J)}{\mathrm{Var}(\tilde{I})}\bar{\tilde{I}} - J(x)\right|^2 dx \\
&= -\frac{1}{\sigma^2}|X|\frac{\mathrm{Cov}^2(\tilde{I}, J)}{\mathrm{Var}(\tilde{I})} + \frac{1}{2\sigma^2}|X|\,\mathrm{Var}(J) \\
&= -\frac{1}{\sigma^2}|X|\,\mathrm{Var}(J)\,\mathrm{NCC}^2(\tilde{I}, J) + \frac{1}{2\sigma^2}|X|\,\mathrm{Var}(J)
\end{aligned}$$

Up to constants that do not depend on the deformation, minimizing sum of square error with an unknown affine intensity transformation is equivalent to maximizing normalized cross correlation (NCC) squared.

□

The Missing Data Problem

We approach the problem of missing or censored data using the Expectation Maximization algorithm [31]. We discretize our problem by associating to the images the lattice of sites $\{\Delta x_i, i = 1, \ldots, S\}$ which are a disjoint partition (voxels) $X = \cup_{i=1}^{S} \Delta x_i$. Each atlas type $A \in \mathcal{A} = \{a_1, a_2, \ldots, \}$ includes a deformation given by $\varphi_A = \int_0^1 v_t^A \circ \phi_t^A dt + id$, so the discrete values can be written as

$$\tilde{I}_i^A = \int_{\Delta x_i} I^A \circ \varphi_A^{-1}(y)dy \ .$$

The observable random field $J_i, i = 1, \ldots, S$, is conditionally Gaussian with constant variances σ_A^2 and mean fields $\tilde{I}_i^A$ determined by the atlas type A. In practice, we choose I^{a_1} as our atlas image, and I^{a_i} for $i \neq 1$ to be constant images. Since they are constant, we need not optimize over the deformations $v_t^{a_i}$ for $i > 0$.

Associate to the measured *incomplete-data* $\mathbf{Y} = \{J_i, i = 1, 2, \ldots, S\}$ the *complete-data* $\mathbf{X} = \{(J_i, A_i), i = 1, \ldots, S\}$, an augmentation with labels determining the atlas types $A_i \in \mathcal{A} = \{a_1, a_2, \ldots, \}$. Model J_i as a Gaussian random variable with mean $\tilde{I}_i^{A_i}$ and variance $\sigma_{A_i}^2$. The complete-data penalized log-likelihood becomes:

$$\log f(\mathbf{X}; v) = \sum_{a \in \mathcal{A}} -\frac{1}{2\sigma^2}\int_0^1 \|v_t^a\|_V^2 dt - \frac{1}{2\sigma_a^2}\sum_{i=1}^{S} \delta_a(A_i)|J_i - \tilde{I}_i^a|^2 \ .$$

The Kronecker-delta $\delta_a(\cdot)$ is 1 when the argument is a, and zero otherwise. The Expectation step (E-step) replaces these functions with their expected value, a posterior probability $\pi_i(a)$ at each voxel.

Theorem 2. *The Expectation Maximization algorithm performs:*

$$v^{a,old} \leftarrow v^{a,new}, \text{ with } \dot{\varphi}_t^{a,old} = v_t^{a,old} \circ \varphi_t^{a,old}, t \in [0,1], \varphi_0^a = id$$
$$\tilde{I}^{a,old} = I^a \circ \varphi^{old-1}, a \in \mathcal{A},$$

E-step: $E\left(log f(\mathbf{X}; v) | \mathbf{Y}, v^{a,old}\right)$

$$\pi_i^{new}(a) = \frac{\frac{1}{\sqrt{2\pi\sigma_a^2}} \exp(-\frac{1}{2\sigma_a^2}|J_i - \tilde{I}_i^{a,old}|^2)}{\sum_{a' \in \mathcal{A}} \frac{1}{\sqrt{2\pi\sigma_{a'}^2}} \exp(-\frac{1}{2\sigma_{a'}^2}|J_i - \tilde{I}_i^{a',old}|^2)} \ , \ a \in \mathcal{A} \ ;$$

M-Step: $\arg\max E\left(\log f(\mathbf{X}; v) | \mathbf{Y}, v^{old}\right)$

$$v^{a,new} = \arg\max_{v^a} -\frac{1}{2\sigma^2} \int_0^1 \|v_t^a\|_V^2 dt - \frac{1}{2\sigma_a^2} \sum_i \pi_i^{new}(a) |J_i - \tilde{I}_i^{a,old}|^2 \ .$$

Iterations $v^{a,new}$ with $\varphi^{a,new} = \int_0^1 v^{a,new} \circ \phi^{a,new} dt + id$ increase in incomplete-data log-likelihood.

The M step updates v_t which is just weighted LDDMM with a weighted L2 cost. Equation (3b) is updated to include the posterior weights (derived for example in [35]), and (3a) is updated as weighted least squares, giving

$$Av_t = \frac{1}{\sigma^2} D\left(F_\theta(I \circ \varphi_t^{-1})\right)^T \left(J(\varphi_{1t}^{-1}) - F_\theta(I \circ \varphi_t^{-1})\right) |D\varphi_{1t}^{-1}| \pi_{\varphi_{1t}^{-1}(\cdot)}(a) \tag{5a}$$

$$\int_X J \cdot \frac{d}{d\theta} F_\theta(I \circ \varphi_1^{-1}) \pi_\cdot(a) dx = \int_X F_\theta(I \circ \varphi_1^{-1}) \cdot \frac{d}{d\theta} F_\theta(I \circ \varphi_1^{-1}) \pi_\cdot(a) dx \tag{5b}$$

In (5a) and (5b), $\pi_\cdot(a)$ are considered functions of space rather than voxel indices.

Post mortem imaging

Preparation and scanning of brain tissue was performed by the neuropathological team at the Johns Hopkins Brain Resource Center (BRC) and the laboratory of Dr. Susumu Mori. The specimen was a 1290 gram 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) [36]; CERAD neuritic plaque score B [37]; neurofibrillary Braak stage VI/VI [38]; subacute infarcts frontal, temporal, and basal ganglia; old infarct of pons; with clinical-pathological 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 orientation 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 [39, 40] 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 percent 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 hours.

The MRI scanning procedure resulted in several distinct images which must be themselves 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. Shown in Fig. 1 left is an example of our images and their alignment. 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 in Fig 1 right. The superimposed lines correspond to the rostral-caudal pages of the Mai-Paxinos-Voss atlas (z axis), with their 1cm grid lines along the x-y axes of each page. Our entorhinal cortex definition (orange) includes the medial bank of the collateral sulcus. This sulcal region [41], also referred to as the trans entorhinal cortex, corresponds to the earliest location of AD pathology accumulation visible at autopsy [42]. Atrophy in this region has been detected at the population level in subjects with mild cognitive impairment [43] before other changes are visible [44].

Fig 1. High-field MTL volume from MR. Left: MR images of tissue blocks aligned in 3D. Right: surface rendering of manual segmentations with Mai Atlas coordinate system superimposed.

After the tissue blocks underwent high field imaging, they were sectioned for histological examination. The tissue was sectioned sectioned at 200μm intervals: 20 slices of 10-μm thickness and 5 slices of 40-μm 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 Aβ (mab 6E10) and tau (PHF1). Several examples of our tau, amyloid, and myelin stained sections are shown in the result section, Fig. 4 to 9. 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 will be used for quantitative cell and neuron counts and density studies of dendritic and synaptic markers.

Image mapping experiments

Below we present several 2D experiments to demonstrate the applicability and verify the validity of our method. In each case, rigid and deformable registration is performed simultaneously using gradient descent, and deformable registration is implemented using LDDMM (3a) or weighted LDDMM (5a). For registration purposes, histology images are downsampled by averaging over a 32x32 window

Our 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 post mortem MRI and 2D histology slices.

1. 3D deformation
2. 3D rigid positioning
3. Estimation of slice spacing (z axis scale)
4. Estimation of pixel size (x, y axis uniform scale)
5. 2D rigid motion (for each slice and stain)
6. Cubic intensity transformation (for each slice and stain)

Rigid, scale, and deformation parameters are all jointly optimized using gradient descent. Deformable registration uses a relatively small gradient descent step size allow 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 [45]. Input data is 56x56 regions of interest. The network uses 3 convolution layers with a 5x5 kernels, max pool downsampling by a factor of 2x2, followed by a fully connected layer and a cross entropy loss function. The center pixel of each region is classified as belonging to one of 3 classes "tau tangles", "other tissue", or "background". 2391 training examples were used, with 8.3% positive, and neural network weights were trained using the Adam optimizer [46]. Every pixel in our histology data was classified using a sliding window approach. 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 structures (which may be affected by missing tissue), and the fraction of this area covered by tau tangles.

Results

Mapping simulated images with artifact and missing data

To demonstrate the method we start with simulated images. Fig 2 shows the atlas (top left panel) and target (top center panel). Contrast is chosen so that the atlas is 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 standard deviation 2/3 pixels over a 5x5 pixel window.

Fig 2. Simulation results Top row: atlas (left) and target (right). Bottom row: Transformed atlas to target; estimated final weights; estimated non-monotonic polynomial F_θ.

The bottom left panel of Fig. 2 shows that a cubic intensity transformation (right panel) is sufficient to permute the order of gray and white mater, allowing for accurate matching of the cortical boundary. The bottom center panel 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 missing tissue in blue. Note that the background is magenta because "atlas image" and "missing tissue" both describe the image intensity equally well.

Figure 3 shows the failed results of existing mapping methods, using a linear contrast transform only (i.e. normalized cross correlation), and a fixed mask (top row) or no mask (bottom row). 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.

Fig 3. Alternative methods Top left shows the results using a fixed binary mask, with center showing the estimated monotonic intensity transformation. Bottom left shows the result using no mask, with right showing the estimated intensity transformation.

Mapping histology with missing data and different stains

In Fig. 4 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 nonmonotonic cubic polynomial, allowing for a swapping of brightness from gray matter (1) $\rightarrow$ white mater (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 (1 constant, 3 linear, 6 quadratic, 10 cubic, for each of 3 dimensions). This example illustrates the intensity transformation component of our algorithm in isolation.

Fig 4. Mapping across modality from tau to myelin. A tau stained section through the hippocampus is mapped to a neighboring myelin stained section using a cubic polynomial intensity transform.

In Fig. 5 we map a tau stained slice of medial temporal lobe to its neighbor which has significant missing data due to damaged tissue. This illustrates the missing data component of our algorithm in isolation.

Fig 5. Mapping a tau section with missing tissue. A tau stained section through the hippocampus is mapped to a neighboring damaged section. The weights shown correspond to posterior probability that a given pixel corresponds to our atlas image, with high probability in yellow, probability 0.5 in green, and low probability in blue.

In Fig. 6 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.

Fig 6. Mapping across modality from tau to myelin with missing tissue A a tau stained section through the hippocampus is mapped to a neighboring LFB stained damaged section. The weights shown correspond to posterior probability that a given pixel corresponds to our atlas image, with high probability in yellow, probability 0.5 in green, and low probability in blue.

Mapping histology data to Mai atlas coordinates

Alignment between 3D post mortem MRI, and each of our three 2D stains are shown in Fig 7. For each stain, we show histology images, intensity transformed MRI, and original aligned MRI. 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.

Fig 7. Mapping from MRI to three histology slices Alignment between our 3D MRI and histology slices at 15 locations with three stains is shown. Top to bottom: amyloid, MRI intensity mapped to amyloid, aligned MRI, LFB, MRI intensity mapped to LFB, aligned MRI, Tau, MRI intensity mapped to tau, aligned MRI. .

Example results of our tau detection algorithm are shown in Fig 8. We achieve an accuracy of 0.996 on a test set of 500 examples. Detection results are shown at 3 different scales, with individual detections at the micron level, and densities at the millimeter level.

Fig 8. Tau detection results Tau stained histology sections (top) with the result of our tangle detection (bottom, red) at 3 different scales, zooming out from single tangles (left) to the entire medial temporal lobe (right).

Figure 9 shows several sections of the Mai-atlas along the rostral to caudal axis in millimeters. In the same 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 (as in ANTs SyN [47]), and a weighted average of images was computed from the flow: $(1-p)I \circ \varphi_t^{-1} + pJ \circ \varphi_{1t}^{-1}$.

Fig 9. Mapping MRI sections to Mai-Paxinos sections. High-field atlas section at 9.30 mm (top row), 17.20 mm (middle row) and 23.90 mm (bottom row) along caudal-rostral axis of histology sections in Mai-Paxinos coordinates.

Finally, Fig. 10 shows our estimated area covered by tau tangles on each page of the Mai atlas for several structures (entorhinal cortex, subiculum, CA fields). We observe a trend of decreasing tangle concentration in the rostral to caudal direction, which will need to be verified for reproducibility as more specimens become available.

Fig 10. Tau tangles on each page of Mai atlas for several structures Left shows the total area of tau tangleps detected within several anatomical structures (ERC: entorhinal cortex, SUB: subiculum, CA: Cornu Ammonis). Center shows the area of these structures, and right shows the fraction of this area covered by tau tangles.

Discussion

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, (iii) the addition of Gaussian white noise. Here (i) describes the object being imaged, and (ii-iii) describe the imaging process, reflecting the distinction made by Shannon between source and channel. This model allows us to formulate multi modality image matching as a penalized likelihood problem, rather than simply the maximization of an ad hoc 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 alternatives. For 2D histology we demonstrated the effectiveness of each of our two contributions, contrast mapping and tissue classification, in isolation and simultaneously. 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, $F(i) = E_{J|\tilde{I}=i}[J]$. This transformation results in a cost function with the same set of invariances as mutual information.

A third popular image similarity term, introduced in [47], is local normalized cross correlation. We are currently extending our method to include polynomial contrast transformations where coefficients are smooth functions of space. As in local normalized cross correlation, this will allow accommodation of image nonuniformity due to magnetic field inhomogeneities or coil sensitivity in MRI, or variable illumination optical imagery.

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 variance of image noise, background noise, and artifact noise. These must be chosen carefully to reflect physical characteristics of our imaging model. Further, as is typical of Expectation Maximization algorithms, optimization in our setting can be slow and sensitive to initialization. The choice of polynomials here to describe intensity changes was for simplicity only, linear combinations of any basis functions can be included in the same formulation. The estimation of polynomial coefficients can be unstable, and future work will involve investigating alternative bases.

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 [48], 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. We are continuing to acquire post mortem samples, and the single subject results presented here will be augmented over the next several years, enabling a more detailed examination of tau distribution. This work advances the field of brain mapping in two important ways. First it moves away from ad hoc image similarities, and toward statistical models of image formation. While our work used a simple white noise model, this framework has the potential to connect with imaging physics and benefit from known properties of imaging systems, such as their signal transfer and noise performance. 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 [19, 20] and iDISCO [49] images in mouse and rat [50], serially sectioned mouse as part of the BRAIN Initiative Cell Census Network [51], and revisiting older datasets where images were excluded due to artifacts or damaged tissue.

References

1. Mori S, Oishi K, Faria AV, Miller MI. "Atlas-Based Neuroinformatics via MRI: Harnessing Information from Past Clinical Cases and Quantitative Image Analysis for Patient Care". Annual Review of Biomedical Engineering. 2013;15:71–92. doi:10.1146/annurev-bioeng-071812-152335.

2. Miller M, Faria A, Oishi K, S M. High-throughput neuro-imaging informatics. Frontiers in Neuroinformatics;31(7). doi:doi:10.3389/fninf.2013.00031.

3. Talairach J, Szikla G. Application of stereotactic concepts to the surgery of epilepsy. In: Advances in Stereotactic and Functional Neurosurgery 4. Springer; 1980. p. 35–54.

4. Evans AC, Collins DL, Mills S, Brown E, Kelly R, Peters TM. 3D statistical neuroanatomical models from 305 MRI volumes. In: Nuclear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record. IEEE; 1993. p. 1813–1817.

5. Mori S, Oishi K, Faria A, Zijl PCM. MRI Atlas of Human White Matter. Elsevier; 2005.

6. Chuang N, Mori S, Yamamoto A, Jiang H, Ye X, Xu X, et al. An MRI-based atlas and database of the developing mouse brain. Neuroimage. 2011;54(1):80–89.

7. Sunkin SM, Ng L, Lau C, andTerri L Gilbert TD, Thompson CL, Hawrylycz M, et al. Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. Nucleic Acids Res. 2013;(41(Database issue)).

8. Miller MI, Christensen GE, Amit Y, Grenander U. Mathematical textbook of deformable neuroanatomies. Proceedings of the National Academy of Sciences. 1993;90(24):11944–11948. doi:10.1073/pnas.90.24.11944.

9. Grenander U, Miller MI. Computational anatomy: An emerging discipline. Quarterly of Applied Mathematics. 1998;56(4):617–694.

10. Toga A, Thompson P. Maps of the Brain. Anat Rec. 2001;265(2):37–53.

11. Thompson PM, Toga AW. A framework for computational anatomy. Computing and Visualization in Science. 2002;5(1):13–34. doi:10.1007/s00791-002-0084-6.

12. Ashburner J, Friston KJ. Computational anatomy. In: Friston KJ, Ashburner J, Kiebel SJ, Nichols TE, Penny WD, editors. Statistical Parametric Mapping The Analysis of Functional Brain Images. Academic Press; 2007. p. 49–100.

13. Ashburner J. Computational Anatomy with the SPM Software. Magnetic Resonance Imaging. 2009;27:1163–1174.

14. Durrleman S, Pennec X, Trouvé A, Ayache N. A Forward Model to Build Unbiased Atlases from Curves and Surfaces. Proc of the International Workshop on the Mathematical Foundations of Computational Anatomy MFCA2008. 2008;.

15. Vialard Fx, Risser L, Holm D, Rueckert D. Diffeomorphic Atlas Estimation using Kärcher Mean and Geodesic Shooting on Volumetric Images Kärcher mean. In: Medical image understanding and analysis; 2011.

16. Zhang Y, Zhang J, Oishi K, Faria AV, Jiang H, Li X, et al. Atlas-guided tract reconstruction for automated and comprehensive examination of the white matter anatomy. NeuroImage. 2010;52:1289–1301.

17. Zhang DP, Risser L, Vialard FX, Edwards P, Metz C, Neefjes L, et al. Coronary Motion Estimation from CTA Using Probability Atlas and Diffeomorphic Registration. Medical Imaging and Augmented Reality. 2010;6326:78–87.

18. Ardekani S, Jain A, Jain S, Abraham TP, Abraham MR, Zimmerman S, et al. Matching Sparse Sets of Cardiac Image Cross-Sections Using Large Deformation Diffeomorphic Metric Mapping Algorithm. In: Camara O, Konukoglu E, Pop M, Rhode K, Sermesant M, Young A, editors. Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges. No. 7085 in Lecture Notes in Computer Science. Springer Berlin Heidelberg; 2012. p. 234–243.

19. Chung K, Wallace J, Kim SY, Kalyanasundaram S, Andalman AS, Davidson TJ, et al. Structural and molecular interrogation of intact biological systems. Nature. 2013;497(7449):332.

20. Epp JR, Niibori Y, Hsiang HLL, Mercaldo V, Deisseroth K, Josselyn SA, et al. Optimization of CLARITY for clearing whole brain and other intact organs. eneuro. 2015; p. ENEURO–0022.

21. S Kutten K, Charon N, Miller M, Ratnanather J, Matelsky J, D Baden A, et al. A Large Deformation Diffeomorphic Approach to Registration of CLARITY Images via Mutual Information; 2017. p. 275–282.

22. Avants BB, Grossman M, Gee JC. 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; 2006. p. 50–57.

23. Wu J, Tang X. Fast Diffeomorphic Image Registration via GPU-based Parallel Computing: An Investigation of the Matching Cost Function. In: Proceedings of SPIE Medical Imaging (SPIE-MI); 2018.

24. Pluim JPW, Maintz JBA, Viergever MA. Mutual-information-based registration of medical images: a survey. IEEE Transactions on Medical Imaging. 2003;22(8):986–1004. doi:10.1109/TMI.2003.815867.

25. Yoo TS, Ackerman MJ, Lorensen WE, Schroeder W, Chalana V, Aylward S, et al. Engineering and algorithm design for an image processing API: a technical report on ITK-the insight toolkit. Studies in health technology and informatics. 2002; p. 586–592.

26. Tang X, Mori S, Miller M. Segmentation via multi-atlas LDDMM. MICCAI 2012 Workshop on Multi-Atlas Labeling. 2012; p. 123–133.

27. Tang X, Oishi K, Faria A, Hillis A, MS A, Mori S, et al. Bayesian Parameter Estimation and Segmentation in the Multi-Atlas Random Orbit Model. PLoS ONE. 2013;8(6).

28. Grenander M, Miller M. Pattern Theory: From Representation to Inference. ISBN 978-0199297061. Oxford Press; 2007.

29. Miller MI, Trouvé A, Younes L. Diffeomorphometry and geodesic positioning systems for human anatomy. Technology. 2014;1:36.

30. Miller MI, Arguillère S, Tward DJ, Younes L. Computational anatomy and diffeomorphometry: A dynamical systems model of neuroanatomy in the soft condensed matter continuum. Wiley Interdisciplinary Reviews: Systems Biology and Medicine;0(0):e1425. doi:10.1002/wsbm.1425.

31. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. Journal of the royal statistical society Series B (methodological). 1977; p. 1–38.

32. Mai JK, Paxinos G. The Human Nervous System. Academic Press; 2011.

33. Grenander U. General Pattern Theory. ISBN 978-0198536710. Oxford Science Publications; 1994.

34. Beg MF, Miller MI, Trouvé A, Younes L. Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms. International Journal of Computer Vision. 2005;61(2):139–157. doi:10.1023/B:VISI.0000043755.93987.aa.

35. Ma J, Miller MI, Trouvé A, Younes L. Bayesian template estimation in computational anatomy. NeuroImage. 2008;42(1):252–261.

36. Hyman BT, Phelps CH, Beach TG, Bigio EH, Cairns NJ, Carrillo MC, et al. National Institute on Aging–Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. Alzheimer's & dementia. 2012;8(1):1–13.

37. Mirra SS, Heyman A, McKeel D, Sumi S, Crain BJ, Brownlee L, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD) Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. Neurology. 1991;41(4):479–479.

38. Braak H, Alafuzoff I, Arzberger T, Kretzschmar H, Del Tredici K. Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. Acta neuropathologica. 2006;112(4):389–404.

39. Mori S, Van Zijl PC. A motion correction scheme by twin-echo navigation for diffusion-weighted magnetic resonance imaging with multiple RF echo acquisition. Magnetic resonance in medicine. 1998;40(4):511–516.

40. Xue R, Sawada M, Goto S, Hurn PD, Traystman RJ, van Zijl PC, et al. Rapid three-dimensional diffusion MRI facilitates the study of acute stroke in mice. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine. 2001;46(1):183–188.

41. Krimer LS, Hyde TM, Herman MM, Saunders RC. The entorhinal cortex: an examination of cyto-and myeloarchitectonic organization in humans. Cerebral cortex (New York, NY: 1991). 1997;7(8):722–731.

42. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta neuropathologica. 1991;82(4):239–259.

43. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. 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 & dementia. 2011;7(3):270–279.

44. Tward DJ, Sicat CS, Brown T, Bakker A, Gallagher M, Albert M, et al. Entorhinal and transentorhinal atrophy in mild cognitive impairment using longitudinal diffeomorphometry. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring. 2017;9:41–50.

45. Tward DJ, Brown T, Patel J, Kageyama Y, Mori S, Troncoso JC, et al. Quantification OF 3D tangle distribution in medial temporal lobe using multimodal image registration and convolutional neural networks. Alzheimer's & Dementia: The Journal of the Alzheimer's Association. 2018;14(7):P57.

46. Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:14126980. 2014;.

47. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Medical image analysis. 2008;12(1):26–41.

48. Adler DH, Pluta J, Kadivar S, Craige C, Gee JC, Avants BB, et al. Histology-derived volumetric annotation of the human hippocampal subfields in postmortem MRI. Neuroimage. 2014;84:505–523.

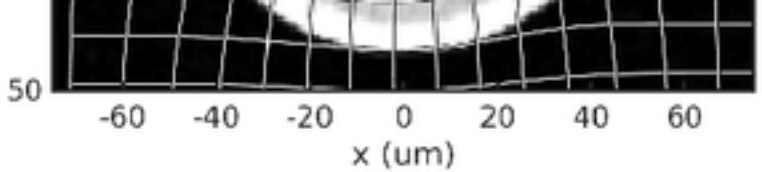
49. Renier N, Wu Z, Simon DJ, Yang J, Ariel P, Tessier-Lavigne M. iDISCO: a simple, rapid method to immunolabel large tissue samples for volume imaging. Cell. 2014;159(4):896–910.

50. Branch AE, Tward DJ, Chandrashekar V, Miller M, Vogelstein JT, Gallagher M. Registration methodology for cleared rodent brain tissue. In: Society for Neuroscience Annual Meeting 2018; 2018. p. 611.12 / LLL39.

51. Lee BC, Tward DJ, Mitra PP, Miller MI. On variational solutions for whole brain serial-section histology using the computational anatomy random orbit model. arXiv preprint arXiv:180203453. 2018;.

52. Towns J, Cockerill T, Dahan M, Foster I, Gaither K, Grimshaw A, et al. XSEDE: accelerating scientific discovery. Computing in Science & Engineering. 2014;16(5):62–74.

Fig. 1 in eps format



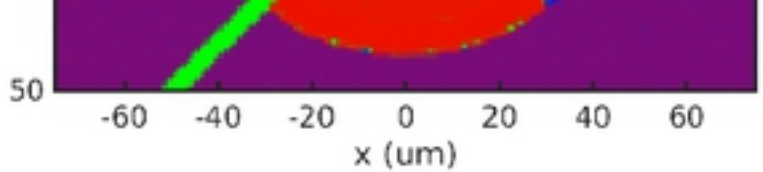


Fig. 2 in eps format

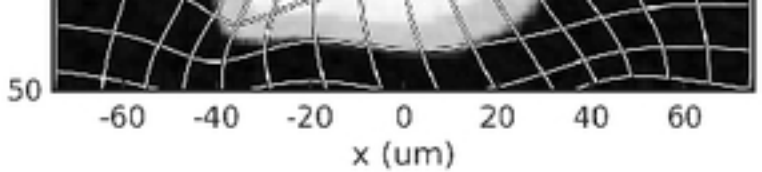


Fig. 3 in eps format

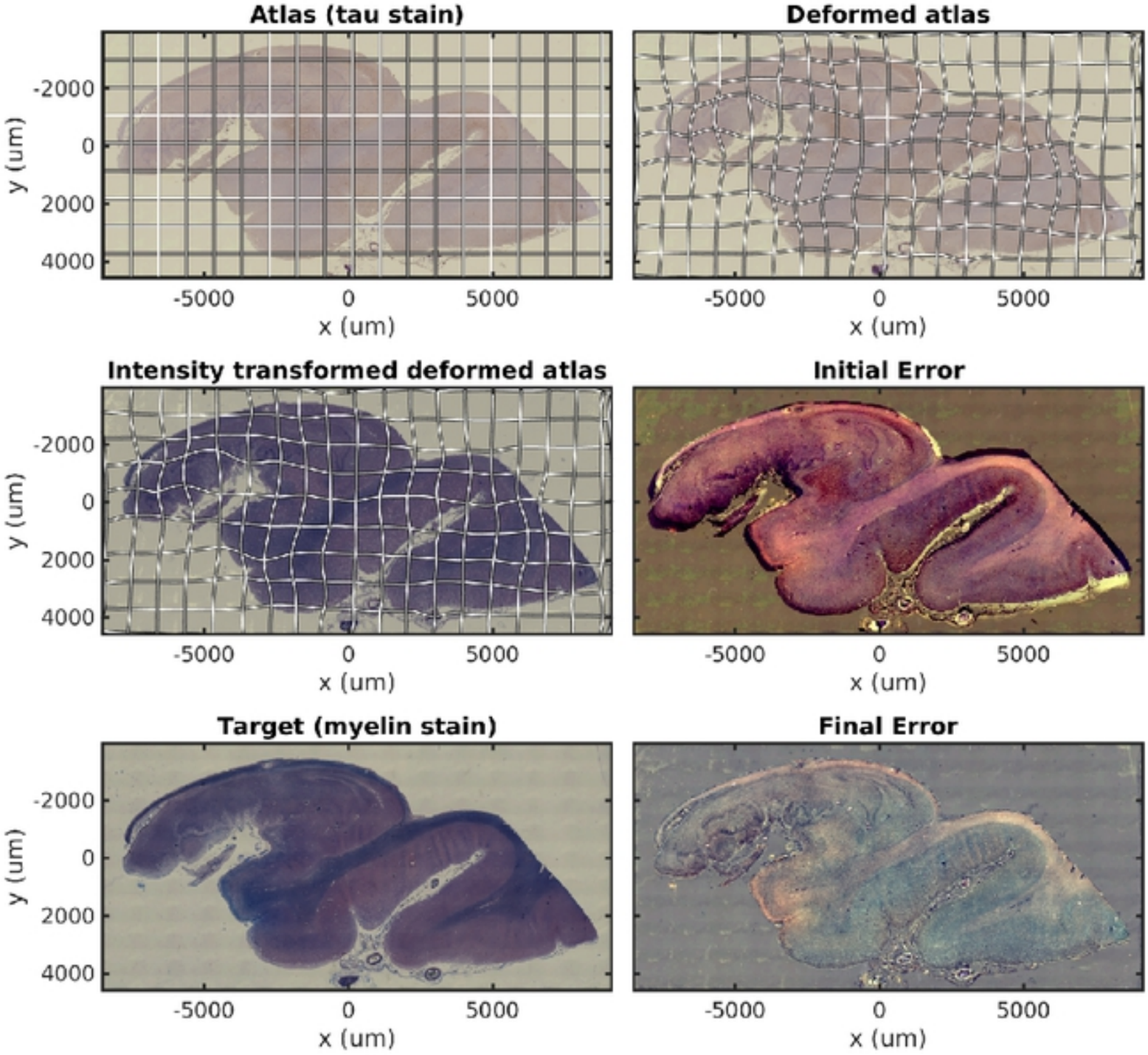


Fig. 4 in eps format

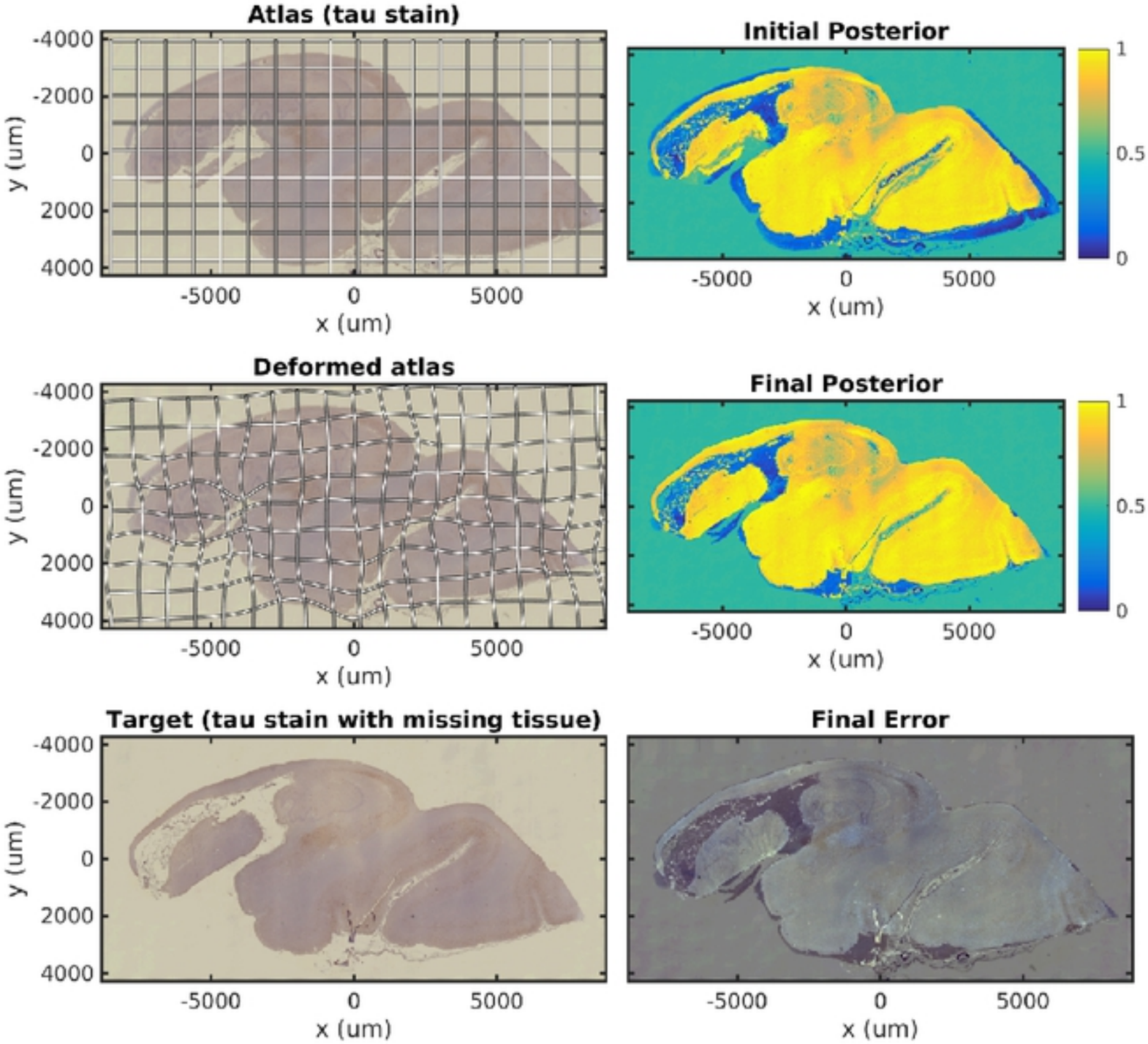


Fig. 5 in eps format

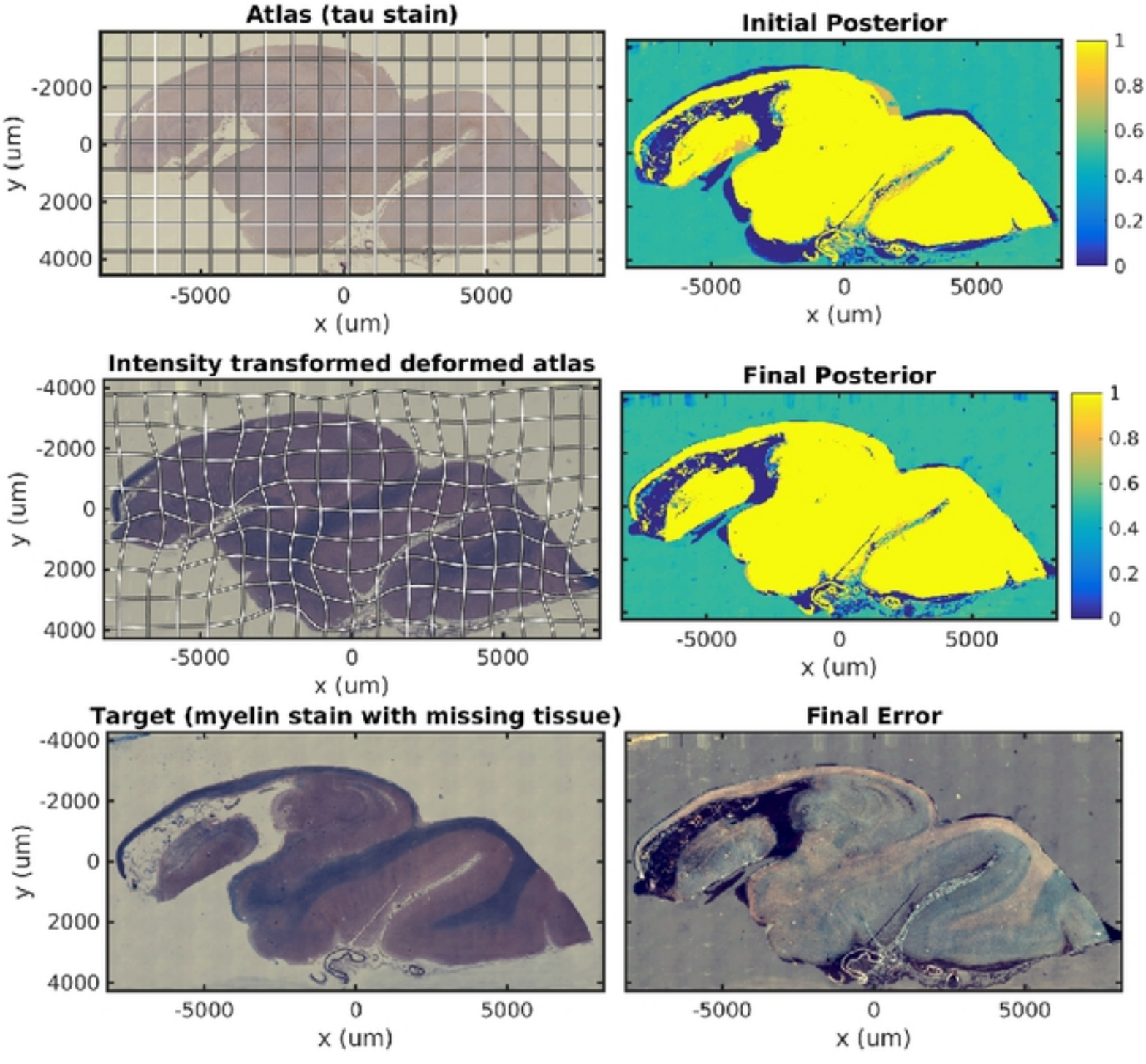


Fig. 6 in eps format

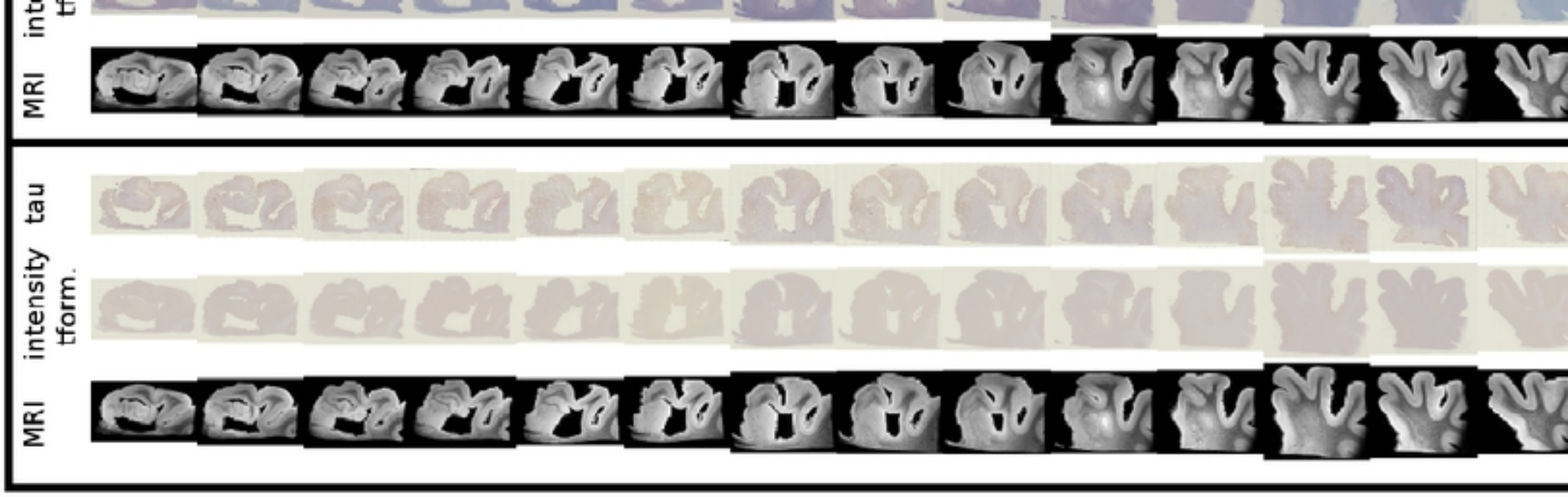


Fig. 7 in eps format

Fig. 8 in eps format

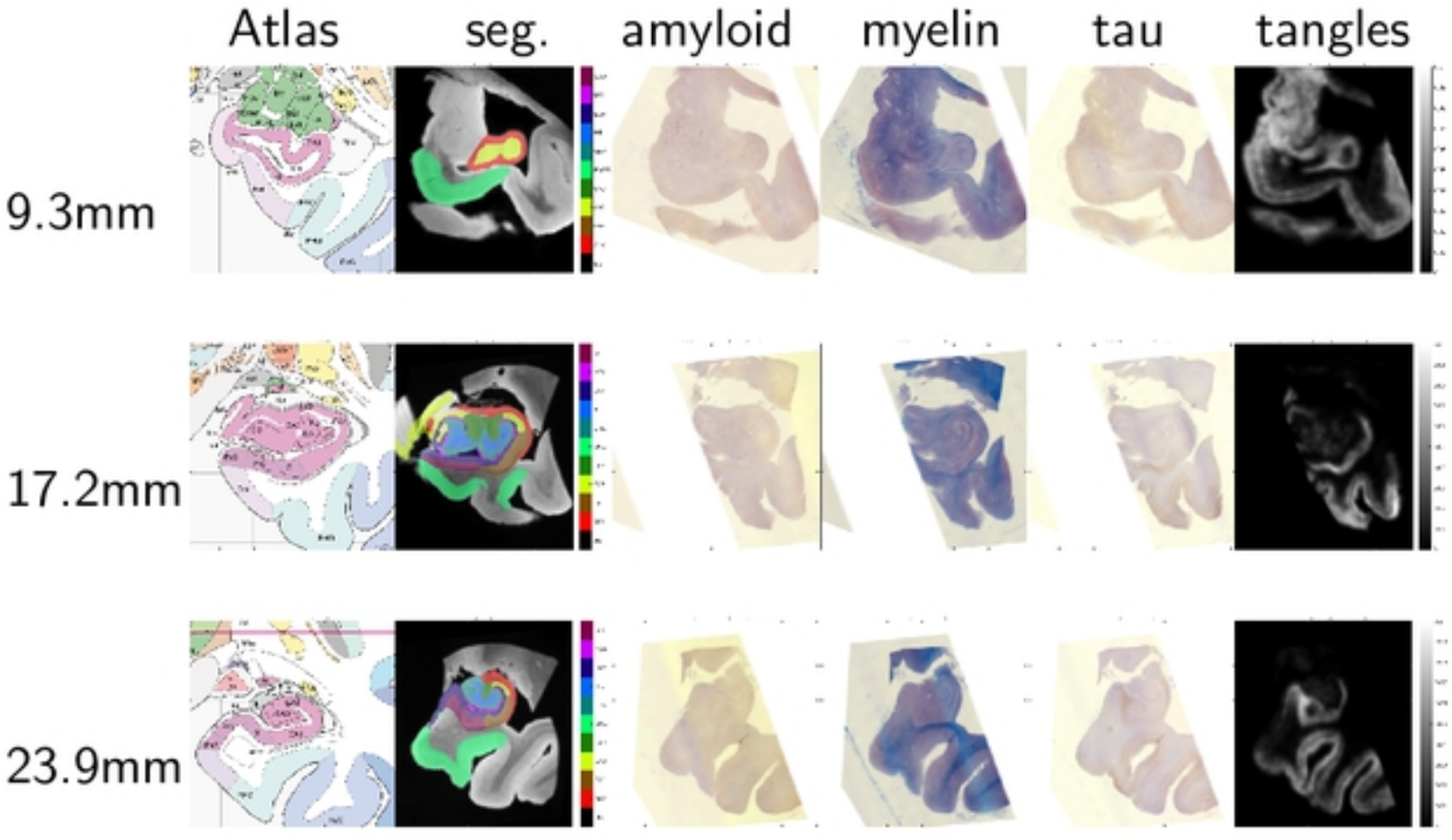


Fig. 9 in eps format

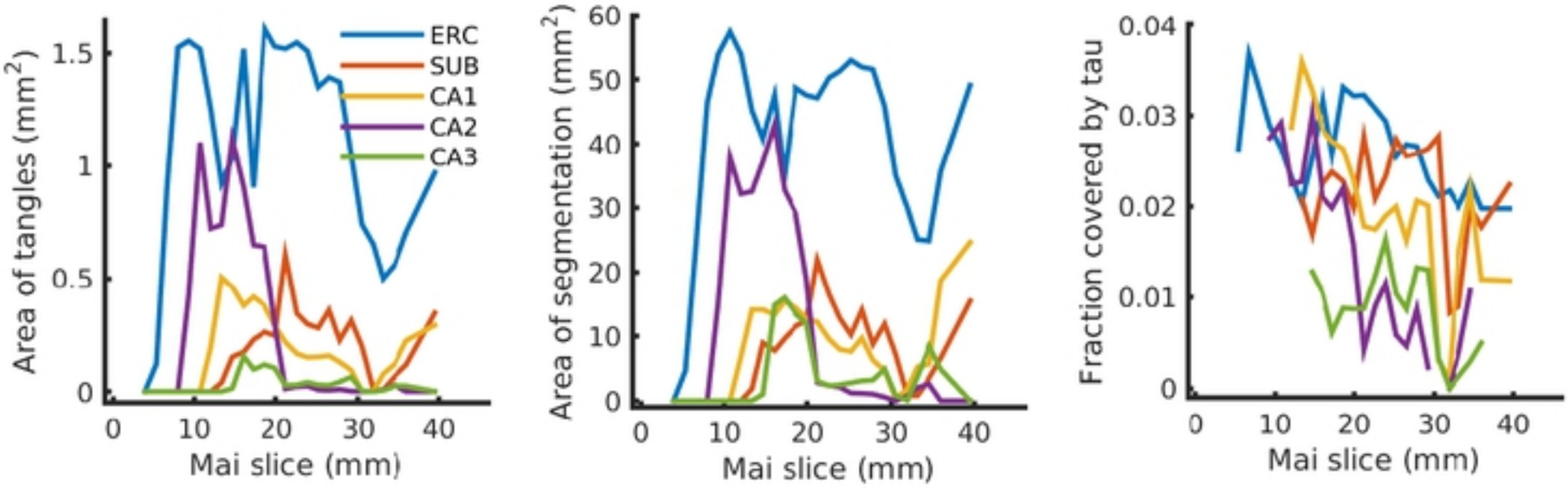


Fig. 10 in eps format