

1

2 Modular strategies for spatial mapping of 3 multi-modal mouse brain data

4 Nicholas J. Tustison¹, Min Chen², Fae N. Kronman³, Jeffrey T. Duda², Clare Gamlin⁴, Mia
5 G. Tustison, Michael Kunst⁴, Rachel Dalley⁴, Staci Sorenson⁴, Quanxin Wang⁴, Lydia Ng⁴,
6 Yongsoo Kim³, and James C. Gee²

7 ¹Department of Radiology and Medical Imaging, University of Virginia, Charlottesville, VA

8 ²Department of Radiology, University of Pennsylvania, Philadelphia, PA

9 ³Department of Neural and Behavioral Sciences, Penn State University, Hershey, PA

10 ⁴Allen Institute for Brain Science, Seattle, WA

11

12 Corresponding authors:

13

14 Nicholas J. Tustison, DSc

15 Department of Radiology and Medical Imaging

16 University of Virginia

17 ntustison@virginia.edu

18

19 James C. Gee, PhD

20 Department of Radiology

21 University of Pennsylvania

22 gee@upenn.edu

²³ **Abstract**

²⁴ Large-scale efforts by the BRAIN Initiative Cell Census Network (BICCN) are generating a
²⁵ comprehensive reference atlas of cell types in the mouse brain. A key challenge in this effort
²⁶ is mapping diverse datasets, acquired with varied imaging, tissue processing, and profiling
²⁷ methods, into shared coordinate frameworks. Here, we present modular mapping pipelines
²⁸ developed using the Advanced Normalization Tools Ecosystem (ANTsX) to align MERFISH
²⁹ spatial transcriptomics and high-resolution fMOST morphology data to the Allen Common
³⁰ Coordinate Framework (CCFv3), and developmental MRI and LSFM data to the Devel-
³¹ opmental CCF (DevCCF). Simultaneously, we introduce two novel methods: 1) a velocity
³² field-based approach for continuous interpolation across developmental timepoints, and 2)
³³ a deep learning framework for automated brain parcellation using minimally annotated and
³⁴ publicly available data. All workflows are open-source and reproducible. We also provide
³⁵ general guidance for selecting appropriate strategies across modalities, enabling researchers
³⁶ to adapt these tools to new data.

³⁷ 1 Introduction

³⁸ Over the past decade, there have been significant advancements in mesoscopic single-cell
³⁹ analysis of the mouse brain. It is now possible to track single neurons¹, observe whole-
⁴⁰ brain developmental changes at cellular resolution², associate brain regions with genetic
⁴¹ composition³, and locally characterize neural connectivity⁴. These scientific achievements
⁴² have been propelled by high-resolution profiling and imaging techniques that enable sub-
⁴³ micron, multimodal, 3D characterizations of whole mouse brains. Among these are micro-
⁴⁴ optical sectioning tomography^{5,6}, tissue clearing methods^{1,7}, spatial transcriptomics^{8,9}, and
⁴⁵ single-cell genomic profiling¹⁰, each offering expanded specificity and resolution for cell-level
⁴⁶ brain analysis.

⁴⁷ Recent efforts by the NIH BRAIN Initiative have mobilized large-scale international collabora-
⁴⁸ tions to create a comprehensive reference database of mouse brain structure and function.
⁴⁹ The BRAIN Initiative Cell Census Network has aggregated over 40 multimodal datasets from
⁵⁰ more than 30 research groups¹¹, many of which are registered to standardized anatomical co-
⁵¹ ordinate systems to support integrated analysis. Among the most widely used of these frame-
⁵² works is the Allen Mouse Brain Common Coordinate Framework (CCFv3)¹². Other CCFs
⁵³ include modality-specific references^{13–15} and developmental atlases^{16,17} that track structural
⁵⁴ change across time.

⁵⁵ 1.1 Mouse brain mapping challenges

⁵⁶ Robust mapping of cell type data into CCFs is essential for integrative analysis of morphology,
⁵⁷ connectivity, and molecular identity. However, each modality poses unique challenges. For
⁵⁸ example, differences in tissue processing, imaging protocols, and anatomical completeness
⁵⁹ often introduce artifacts such as distortion, tearing, holes, and signal dropout^{18–23}. Inten-
⁶⁰ sity differences and partial representations of anatomy can further complicate alignment.
⁶¹ Also, while alternative strategies for mapping single-cell spatial transcriptomic data exist
⁶² (e.g., gene expression-based models such as Tangram²⁴) this work focuses on image-based
⁶³ anatomical alignment to common coordinate frameworks using spatially resolved reference

64 images.

65 Given this diversity specialized strategies are often needed to address the unique, modality-
66 specific challenges.

67 Existing mapping solutions fall into three broad categories. The first includes integrated
68 processing platforms that provide users with mapped datasets (e.g., Allen Brain Cell
69 Atlas²⁵, Brain Architecture Portal²⁶, OpenBrainMap²⁷, and Image and Multi-Morphology
70 Pipeline²⁸). These offer convenience and high-quality curated data, but limited gener-
71 alizability and customization. The second category involves highly specialized pipelines
72 tailored to specific modalities such as histology^{29–31}, magnetic resonance imaging (MRI)^{32–34},
73 microCT^{35,36}, light sheet fluorescence microscopy (LSFM)^{37,38}, fluorescence micro-optical
74 sectioning tomography (fMOST)^{15,39}, and spatial transcriptomics, including multiplexed
75 error-robust fluorescence *in situ* hybridization (MERFISH)^{40–42}. While effective, these
76 solutions often require extensive engineering effort to adapt to new datasets or modal-
77 ities. Finally, general-purpose toolkits such as elastix⁴³, Slicer3D⁴⁴, and the Advanced
78 Normalization Tools Ecosystem (ANTsX)⁴⁵ have all been applied to mouse brain mapping
79 scenarios. These toolkits support modular workflows that can be flexibly composed from
80 reusable components, offering a powerful alternative to rigid, modality-specific solutions.
81 However, their use often requires familiarity with pipeline modules, parameter tuning, and
82 tool-specific conventions which can limit adoption.

83 Building on this third category, we describe a set of modular, ANTsX-based pipelines specif-
84 ically tailored for mapping diverse mouse brain data into standardized anatomical frame-
85 works. These include two new pipelines: a velocity field–based interpolation model that
86 enables continuous transformations across developmental timepoints of the DevCCF, and a
87 template-based deep learning pipeline for whole brain segmentation (i.e., brain extraction)
88 and structural anatomical regional labeling of the brain (i.e., brain parcellation) requiring
89 minimal annotated data. In addition, we include two modular pipelines for aligning MER-
90 FISH and fMOST datasets to the Allen CCFv3. These workflows were adapted and tailored
91 using ANTsX tools to support collaborative efforts within the BICCN and are now made
92 openly available in a reproducible format. To facilitate broader adoption, we also provide

93 general guidance for customizing these strategies across imaging modalities and data types.
94 We first introduce key components of the ANTsX toolkit, which provide a basis for all of the
95 mapping workflows described here, and then detail the specific contributions made in each
96 pipeline.

97 1.2 The Advanced Normalization Tools Ecosystem (ANTsX)

98 The Advanced Normalization Tools Ecosystem (ANTsX) has been used in a number of
99 applications for mapping mouse brain data as part of core processing steps in various
100 workflows^{31,46–49}, particularly its pairwise, intensity-based image registration capabilities⁵⁰
101 and bias field correction⁵¹. Historically, ANTsX development is based on foundational ap-
102 proaches to image mapping^{52–54}, especially in the human brain, with key contributions such
103 as the Symmetric Normalization (SyN) algorithm⁵⁰. It has been independently evaluated
104 in diverse imaging domains including multi-site brain MRI⁵⁵, pulmonary CT⁵⁶, and multi-
105 modal brain tumor registration⁵⁷.

106 Beyond registration, ANTsX provides functionality for template generation⁵⁸, intensity-based
107 segmentation⁵⁹, preprocessing^{51,60}, and deep learning⁴⁵. It has demonstrated strong perfor-
108 mance in consensus labeling⁶¹, brain tumor segmentation⁶², and cardiac motion estimation⁶³.
109 Built on the Insight Toolkit (ITK)⁶⁴, ANTsX benefits from open-source contributions while
110 supporting continued algorithm evaluation and innovation. In the context of mouse brain
111 data, ANTsX provides a robust platform for developing modular pipelines to map diverse
112 imaging modalities into CCFs. These tools span multiple classes of mapping problems:
113 cross-modality image registration, landmark-driven alignment, temporal interpolation across
114 developmental stages, and deep learning-based segmentation. As such, they also serve as il-
115 lustrative case studies for adapting ANTsX tools to other use cases. We describe both shared
116 infrastructure and targeted strategies adapted to the specific challenges of each modality.
117 This paper highlights usage across distinct BICCN projects such as spatial transcriptomic
118 data from MERFISH, structural data from fMOST, and multimodal developmental data
119 from LSFM and MRI.

¹²⁰ **1.3 Novel ANTsX-based open-source contributions**

¹²¹ We introduce two novel contributions to ANTsX developed as part of collaborative efforts
¹²² in creating the Developmental Common Coordinate Framework (DevCCF)¹⁶. First, we
¹²³ present an open-source velocity field-based interpolation framework for continuous map-
¹²⁴ ping across the sampled embryonic and postnatal stages of the DevCCF atlas¹⁶. This
¹²⁵ functionality enables biologically plausible interpolation between timepoints via a time-
¹²⁶ parameterized diffeomorphic velocity model⁶⁵, inspired by previous work⁶⁶. Second, we
¹²⁷ present a deep learning pipeline for structural parcellation of the mouse brain from mul-
¹²⁸ timodal MRI data. This includes two novel components: 1) a template-derived brain ex-
¹²⁹ traction model using augmented data from two ANTsX-derived template datasets^{67,68}, and
¹³⁰ 2) a template-derived parcellation model trained on DevCCF P56 labelings mapped from
¹³¹ the AllenCCFv3. This pipeline demonstrates how ANTsX tools and public resources can be
¹³² leveraged to build robust anatomical segmentation pipelines with minimal annotated data.
¹³³ We independently evaluate this framework using a longitudinal external dataset⁶⁹, demon-
¹³⁴ strating generalizability across specimens and imaging protocols. All components are openly
¹³⁵ available through the R and Python ANTsX packages, with general-purpose functionality
¹³⁶ documented in a reproducible, cross-platform tutorial (<https://tinyurl.com/antsxtutorial>).
¹³⁷ Code specific to this manuscript, including scripts to reproduce the novel contributions
¹³⁸ and all associated evaluations, is provided in a dedicated repository (<https://github.com/>
¹³⁹ [ntustison/ANTsXMouseBrainMapping](https://github.com/ntustison/ANTsXMouseBrainMapping)). Additional tools for mapping spatial transcriptomic
¹⁴⁰ (MERFISH) and structural (fMOST) data to the AllenCCFv3 are separately available at
¹⁴¹ (<https://github.com/dontminchenit/CCFAAlignmentToolkit>).

142 **2 Results**

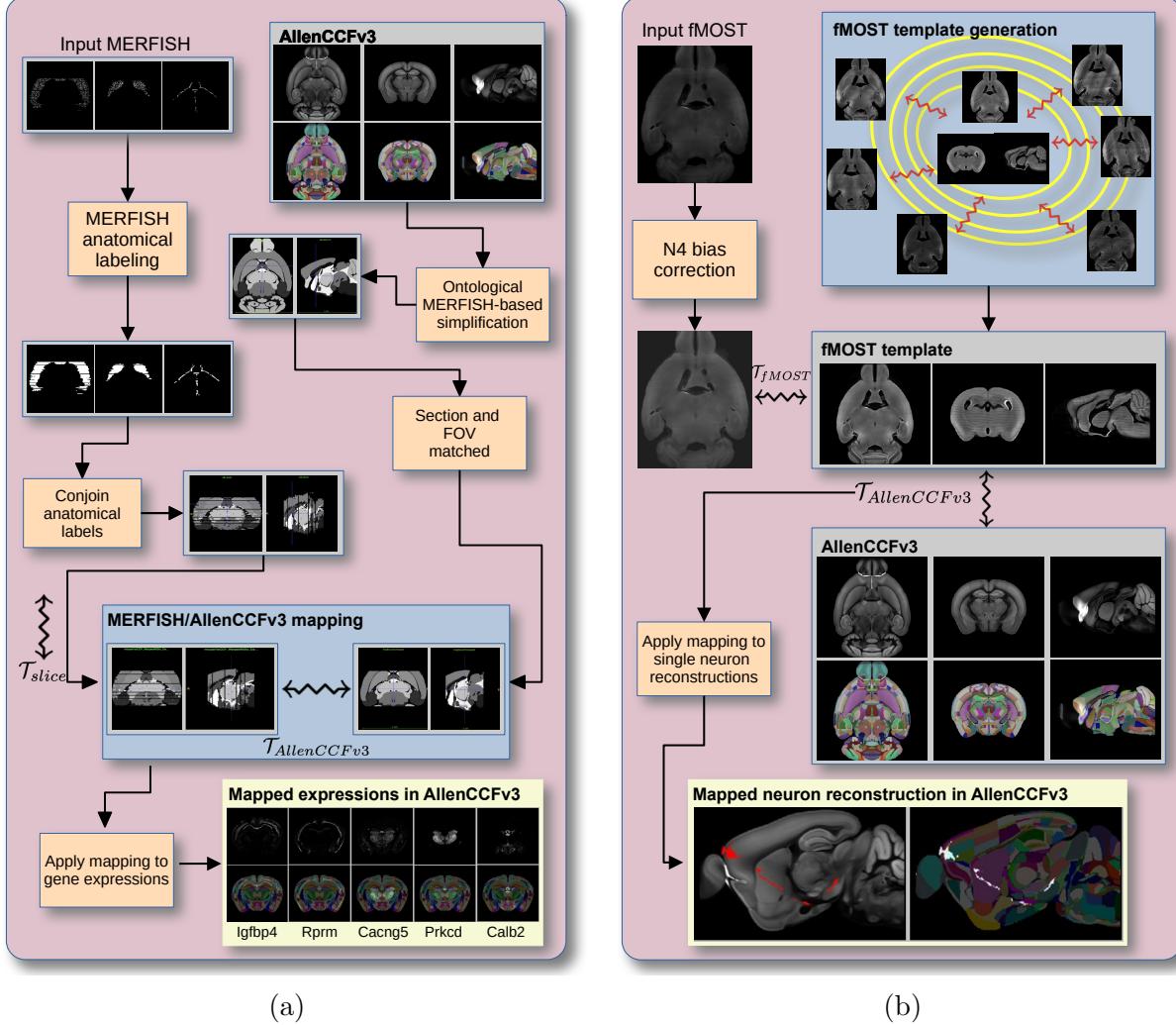


Figure 1: Diagram of the two ANTsX-based pipelines for mapping (a) MERFISH and (b)fMOST data into the space of AllenCCFv3. Each generates the requisite transforms to map individual images to the CCF.

143 **2.1 AllenCCFv3 brain image mapping**

144 **2.1.1 Mapping multiplexed error-robust fluorescence *in situ* hybridization
(MERFISH) data**

146 **Overview.** We developed an ANTsX-based pipeline to map spatial transcriptomic MER-
147 FISH data into the AllenCCFv3 (Figure 1(a)). This approach was used in recent efforts to

¹⁴⁸ create a high-resolution transcriptomic atlas of the mouse brain⁴⁹. The pipeline maps spa-
¹⁴⁹ tial gene expression patterns from MERFISH onto anatomical labels in the AllenCCFv3. It
¹⁵⁰ includes MERFISH-specific preprocessing steps such as section reconstruction, label genera-
¹⁵¹ tion from spatial transcriptomic maps, and anatomical correspondence mapping. Alignment
¹⁵² proceeds in two stages: 1) 3D affine registration and section matching of the AllenCCFv3 to
¹⁵³ the MERFISH data, and 2) linear + deformable 2D section-wise alignment between matched
¹⁵⁴ MERFISH and atlas slices. These transformations are concatenated to produce a complete
¹⁵⁵ mapping from each MERFISH data to AllenCCFv3.

¹⁵⁶ **Data.** MERFISH imaging was performed on cryosectioned brains from C57BL/6 mice using
¹⁵⁷ previously described protocols⁴⁹. Brains were placed into an optimal cutting temperature
¹⁵⁸ (OCT) compound (Sakura FineTek 4583) stored at -80°. The fresh frozen brain was sectioned
¹⁵⁹ at 10 μm on Leica 3050 S cryostats at intervals of 200 μm to evenly cover the brain. A set
¹⁶⁰ of 500 genes was selected to distinguish \sim 5200 transcriptomic clusters. Raw MERSCOPE
¹⁶¹ data were decoded using Vizgen software (v231). Cell segmentation was performed using
¹⁶² Cellpose^{70,71} based on DAPI and PolyT stains which was propagated to adjacent slices across
¹⁶³ z-planes. Each MERFISH cell was assigned a transcriptomic identity by mapping to a
¹⁶⁴ scRNA-seq reference taxonomy.

¹⁶⁵ **Evaluation.** Alignment quality was evaluated iteratively by an expert anatomist, guided
¹⁶⁶ by expected gene-marker correspondences to AllenCCFv3 regions. As previously reported⁴⁹,
¹⁶⁷ further assessment of the alignment showed that, of the 554 terminal regions (gray matter
¹⁶⁸ only in the AllenCCFv3), only seven small subregions did not contain cells from the MER-
¹⁶⁹ FISH dataset post registration: frontal pole, layer 1 (FRP1), FRP2/3, FRP5; accessory
¹⁷⁰ olfactory bulb, glomerular layer (AOBgl); accessory olfactory bulb, granular layer (AOBgr);
¹⁷¹ accessory olfactory bulb, mitral layer (AOBmi); and accessory supraoptic group (ASO).

¹⁷² 2.1.2 Mapping fluorescence micro-optical sectioning tomography (fMOST) data

¹⁷³ **Overview.** We also constructed a pipeline for mapping fMOST images to the AllenCCFv3
¹⁷⁴ using ANTsX (Figure 1(b)). The approach leverages a modality-specific average fMOST
¹⁷⁵ atlas as an intermediate target, adapted from previous work in human and mouse brain

¹⁷⁶ mapping^{12,15,16,58,72–75}. The atlas was constructed from 30 fMOST images selected to cap-
¹⁷⁷ ture representative variability in anatomical shape and image intensity across the population.
¹⁷⁸ Preprocessing includes cubic B-spline downsampling to match the 25 μm isotropic AllenC-
¹⁷⁹ CFv3 resolution, stripe artifact suppression using a 3D notch filter implemented with SciPy’s
¹⁸⁰ frequency-domain filtering tools, and N4 bias field correction⁵¹. A one-time, annotation-
¹⁸¹ driven alignment registers the fMOST atlas to AllenCCFv3 using landmark-based registra-
¹⁸² tion of key structures. This canonical mapping is then reused. New fMOST specimens are
¹⁸³ first aligned to the fMOST atlas using standard intensity-based registration, and the con-
¹⁸⁴ catenated transforms yield full spatial normalization to the AllenCCFv3. This same mapping
¹⁸⁵ can be applied to neuron reconstructions to facilitate population-level analysis of morphology
¹⁸⁶ and spatial distribution.

¹⁸⁷ **Data.** fMOST imaging was performed on 55 mouse brains with sparse transgenic labeling
¹⁸⁸ of neuron populations^{76,77} using the high-throughput fMOST platform^{78,79}. Voxel resolution
¹⁸⁹ was $0.35 \times 0.35 \times 1.0 \mu\text{m}^3$. Two imaging channels were acquired: GFP-labeled neuron mor-
¹⁹⁰ phology (green), and propidium iodide counterstaining for cytoarchitecture (red). Alignment
¹⁹¹ was performed using the red channel for its greater contrast, though multi-channel mapping
¹⁹² is also supported.

¹⁹³ **Evaluation.** The canonical mapping from the fMOST atlas to AllenCCFv3 was eval-
¹⁹⁴ uated using both quantitative and qualitative approaches. Dice similarity coefficients were
¹⁹⁵ computed between corresponding anatomical labels in the fMOST atlas and AllenCCFv3
¹⁹⁶ following registration. These labels were manually annotated or adapted from existing atlas
¹⁹⁷ segmentations. Representative Dice scores included: whole brain (0.99), caudate putamen
¹⁹⁸ (0.97), fimbria (0.91), posterior choroid plexus (0.93), anterior choroid plexus (0.96), optic
¹⁹⁹ chiasm (0.77), and habenular commissure (0.63). In addition to these quantitative assess-
²⁰⁰ ments, each registered fMOST specimen was evaluated qualitatively. An expert anatomist
²⁰¹ reviewed alignment accuracy and confirmed structural correspondence. Neuron reconstruc-
²⁰² tions from individual brains were also transformed into AllenCCFv3 space, and their trajec-
²⁰³ tories were visually inspected to confirm anatomical plausibility and preservation of known
²⁰⁴ projection patterns.

205 **2.2 Continuously mapping the DevCCF developmental trajectory**

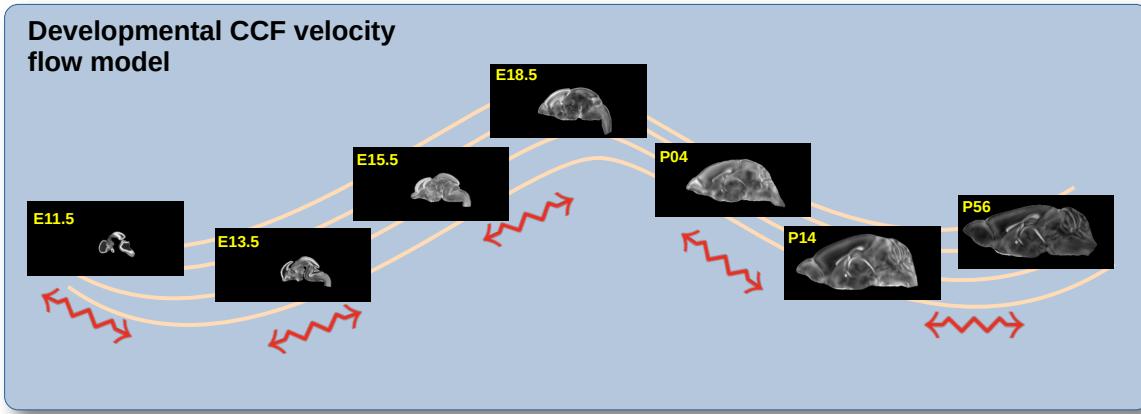


Figure 2: The spatial transformation between any two time points within the continuous DevCCF longitudinal developmental trajectory is available through the use of ANTsX functionality for generating a velocity flow model.

206 The DevCCF is an openly accessible resource for the mouse brain research community⁸⁰,
207 comprising symmetric, multi-modal MRI and LSFM templates generated using the ANTsX
208 framework⁵⁸. It spans key stages of mouse brain development (E11.5, E13.5, E15.5, E18.5,
209 P4, P14, and P56) and includes structural labels defined by a developmental ontology.
210 The DevCCF was constructed in coordination with the AllenCCFv3 to facilitate integra-
211 tion across atlases and data types.

212 Although this collection provides broad developmental coverage, its discrete sampling limits
213 the ability to model continuous transformations across time. To address this, we developed
214 a velocity flow-based modeling approach that enables anatomically plausible, diffeomorphic
215 transformations between any two continuous time points within the DevCCF range. Unlike
216 traditional pairwise interpolation, which requires sequential warping through each interme-
217 diate stage, this model, defined by a time-varying velocity field (i.e., a smooth vector field
218 defined over space and time that governs the continuous deformation of an image domain),
219 allows direct computation of deformations between any two time points in the continuum
220 which improves smoothness and enables flexible spatiotemporal alignment.

221 This functionality is implemented in both ANTsR and ANTsPy (`ants.fit_time_varying_transform_to_`

and integrates seamlessly with existing ANTsX workflows. The velocity field is represented as a 4D ITK image where each voxel stores the x,y,z components of motion at a given time point. Integration of the time-varying velocity field uses uses 4th order Runge-Kutta (`ants.integrate_velocity_field(...)`)⁸¹.

2.2.1 Data

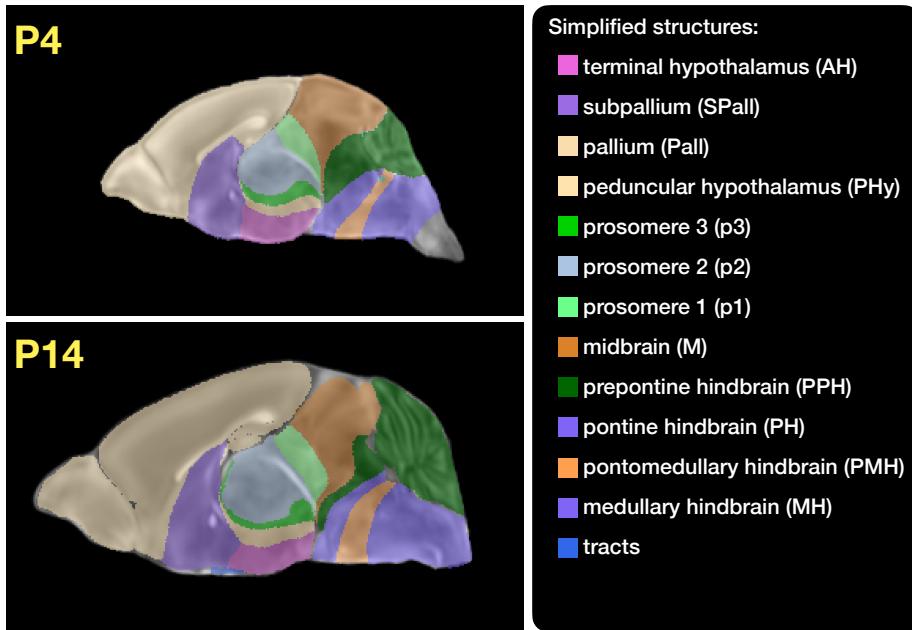


Figure 3: Annotated regions representing common labels across developmental stages, shown for both P4 and P14.

Each DevCCF template includes over 2,500 labeled anatomical regions, with spatial resolutions ranging from 31.5 to 50 μm . For the velocity flow modeling task, we identified a common set of 26 bilateral regions (13 per hemisphere) that were consistently labeled across all timepoints. These regions span major developmental domains including the pallium, subpallium, midbrain, prosomeres, hypothalamus, hindbrain subregions, and key white matter tracts (Figure 3).

Prior to velocity field optimization, all templates were rigidly aligned to the DevCCF P56 template using the centroids of these common label sets. Pairwise correspondence between adjacent timepoints was then computed using ANTsX's multi-metric registration via

236 `ants.registration(...)`. Instead of performing intensity-based multi-label registration di-
237 rectly, we constructed 24 binary label masks per atlas pair (one per structure) and optimized
238 alignment using the mean squares similarity metric with the SyN transform⁵⁰.

239 To generate the point sets for velocity field optimization, we sampled both boundary (con-
240 tour) and interior (region) points from the P56 labels and propagated them to each devel-
241 opmental stage using the learned pairwise transforms. Contours were sampled at 10% of
242 available points and regions at 1%, yielding 173,303 total points per atlas ($N_{contour} = 98,151$;
243 $N_{region} = 75,152$). Boundary points were assigned double weight during optimization to
244 emphasize anatomical boundary correspondence.

245 **2.2.2 Velocity field optimization**

246 The velocity field was optimized using the seven corresponding point sets and their associated
247 weights. The field geometry was defined at [256, 182, 360] with 11 integration points at 50
248 μm resolution, yielding a compressed velocity model of ~ 2 GB. This resolution balanced
249 accuracy and computational tractability while remaining portable. All data and code are
250 publicly available in the accompanying GitHub repository.

251 To normalize temporal spacing, we assigned scalar values in [0, 1] to each template. Given
252 the nonlinear spacing in postnatal development, we applied a logarithmic transform to the
253 raw time values prior to normalization. Within this logarithmic temporal transform, P56
254 was assigned a span of 28 postnatal days to reflect known developmental dynamics (i.e., in
255 terms of modeling the continuous deformation, the morphological changes between Day 28
256 and Day 56 are insignificant). This improved the temporal distribution of integration points
257 (Figure 4, right panel).

258 Optimization was run for a maximum of 200 iterations using a 2020 iMac (3.6 GHz 10-Core
259 Intel Core i9, 64 GB RAM), with each iteration taking ~ 6 minutes. During each iteration,
260 the velocity field was updated across all 11 integration points by computing regularized
261 displacement fields between warped point sets at adjacent time slices. Updates were applied
262 using a step size of $\delta = 0.2$. Convergence was assessed via average displacement error across
263 all points, with final convergence achieved after ~ 125 iterations (Figure 4, left panel).

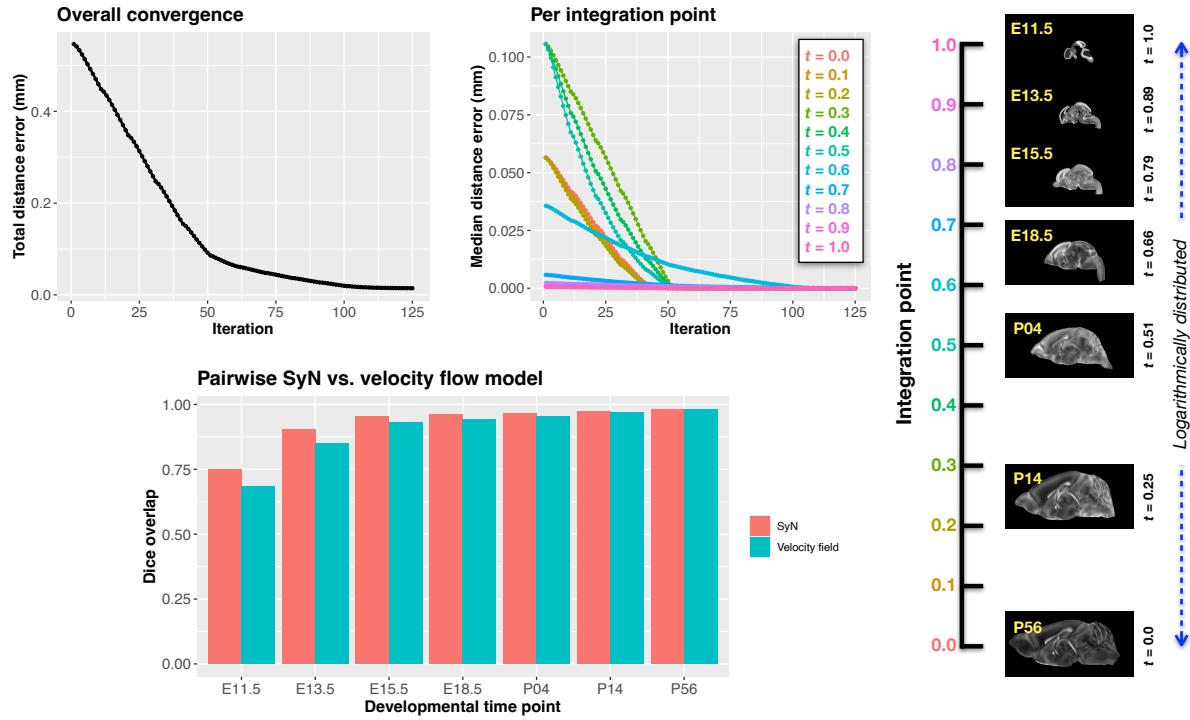


Figure 4: Convergence and evaluation of the velocity flow model across the DevCCF developmental trajectory. (Top left) Total distance error and (top right) per-integration-point error across optimization iterations. where integration spans the full range from embryonic (E11.5) to postnatal (P56) templates (right). (Bottom) Comparison of segmentation overlap accuracy (Dice score) between the velocity flow model and conventional pairwise SyN registration across intermediate DevCCF timepoints. The velocity model achieves comparable accuracy to SyN while also allowing for a smooth continuous deformation across the entire developmental trajectory.

264 Median errors across integration points also trended toward zero, albeit at varying rates.
 265 Using region-based pairwise registration with SyN as a performance upper bound at sampled
 266 timepoints, we find that the velocity flow model achieves comparable accuracy while also
 267 providing a smooth, continuous deformation across the entire developmental trajectory.

268 2.2.3 The velocity flow transformation model

269 Once optimized, the velocity field enables the computation of diffeomorphic transformations
 270 between any pair of continuous time points within the DevCCF developmental range. Fig-
 271 ure 5 illustrates cross-warping between all DevCCF stages using the velocity flow model. In
 272 addition to facilitating flexible alignment between existing templates, the model also sup-

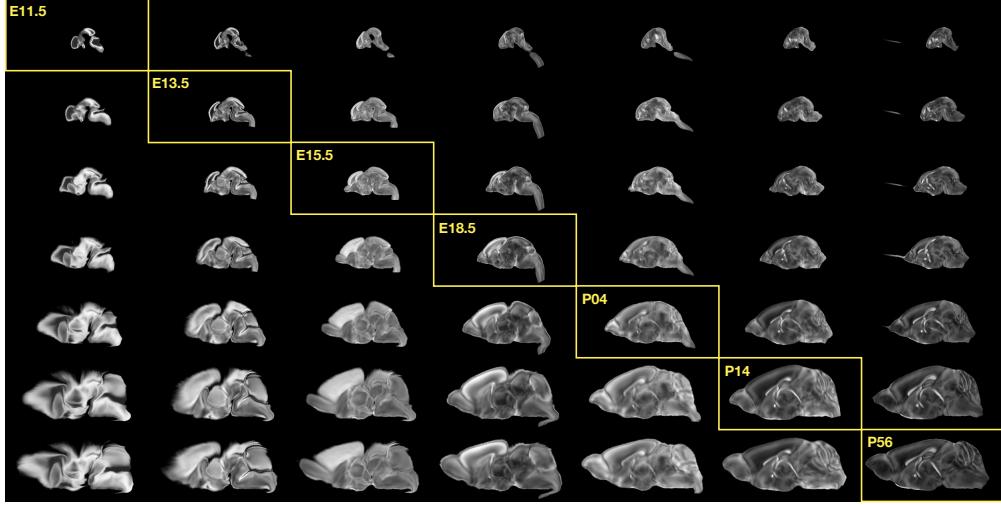


Figure 5: Mid-sagittal visualization of DevCCF templates warped to every other time point. Each row is a reference space; each column is a warped input. Diagonal entries show original templates.

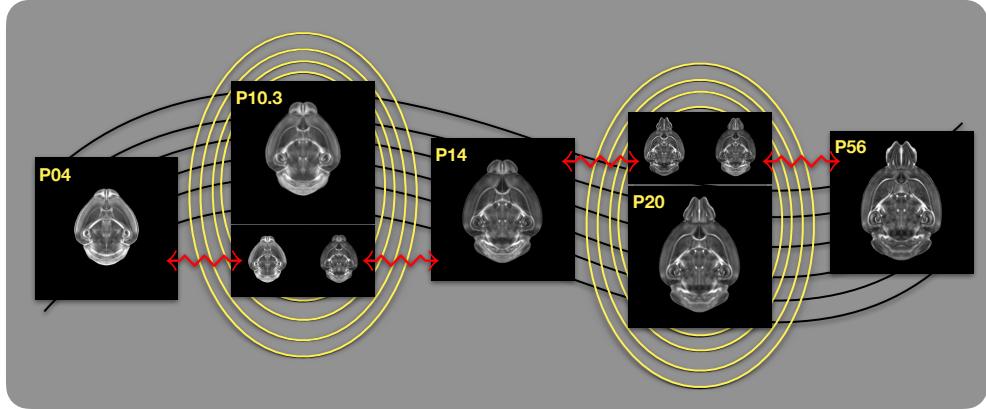


Figure 6: Example of generating “virtual” DevCCF templates at intermediate time points (e.g., P10.3, P20) by warping adjacent stages to a shared time and averaging using ANTsX.

273 ports the synthesis of virtual templates at intermediate, unsampled developmental stages.
 274 As shown in Figure 6, we demonstrate the creation of virtual age templates (e.g., P10.3 and
 275 P20) by warping adjacent developmental atlases to a target timepoint and constructing an
 276 averaged representation using ANTsX’s template-building functionality.
 277 All usage examples, scripts, and supporting data for full reproducibility are publicly available
 278 in the associated codebase.

²⁷⁹ **2.3 Automated structural labeling of the mouse brain**

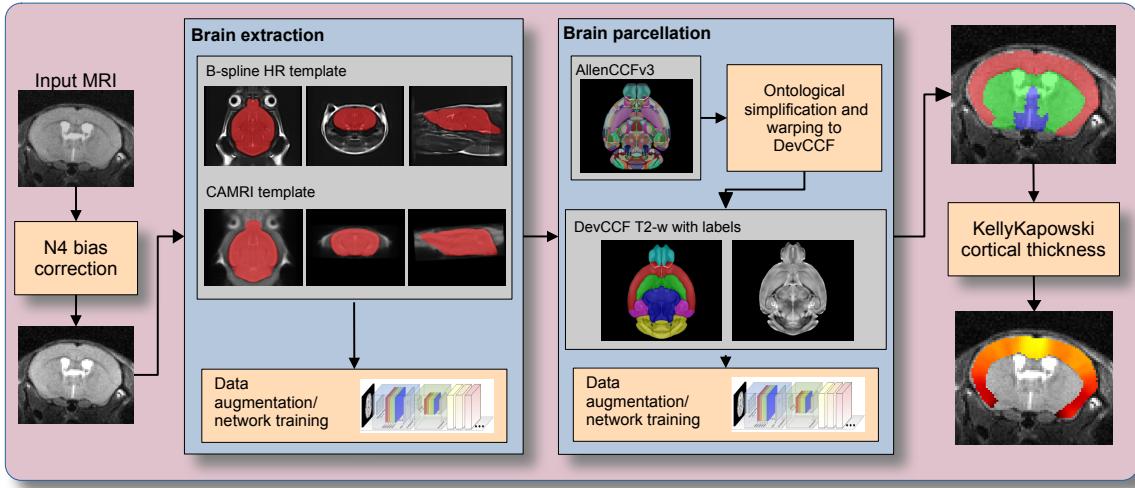


Figure 7: The mouse brain cortical labeling pipeline integrates two deep learning components for brain extraction and anatomical region segmentation. Both networks rely heavily on data augmentation applied to templates constructed from open datasets. The framework also supports further refinement or alternative label sets tailored to specific research needs. Possible applications include voxelwise cortical thickness estimation.

²⁸⁰ Structural labeling strategies for the mouse brain are essential for understanding the organization
²⁸¹ and function of the murine nervous system⁸². By dividing the brain into anatomically
²⁸² or functionally defined regions, researchers can localize biological processes, relate regional
²⁸³ features to behavior, or quantify spatial variation in gene expression patterns^{83,84}. While
²⁸⁴ deep learning techniques have yielded robust segmentation and labeling tools for the hu-
²⁸⁵ man brain (e.g., SynthSeg⁸⁵, ANTsXNet⁴⁵), analogous development for mouse data (e.g.,
²⁸⁶ MEMOS⁸⁶) has been limited. Mouse neuroimaging often presents unique challenges, such
²⁸⁷ as highly anisotropic sampling, that complicate transfer of existing tools. At the same time,
²⁸⁸ high resolution resources like the AllenCCFv3 and DevCCF provide reference label sets that
²⁸⁹ can serve as training data. We demonstrate how ANTsX can be used to construct a full
²⁹⁰ structural labeling pipeline for the mouse brain (Figure 7), including both whole brain seg-
²⁹¹ mentation (i.e., brain extraction) and the subsequent template-based region segmentation.

292 **2.3.1 Template-based mouse brain extraction network**

293 To develop a general-purpose mouse brain extraction model, we constructed whole-head
294 templates from two publicly available T2-weighted datasets. The first dataset, from the
295 Center for Animal MRI (CAMRI) at the University of North Carolina at Chapel Hill⁶⁷,
296 includes 16 isotropic MRI volumes acquired at $0.16 \times 0.16 \times 0.16$ mm³ resolution. The second
297 dataset⁶⁸ comprises 88 specimens acquired in three orthogonal 2D views (coronal, axial,
298 sagittal) at 0.08×0.08 mm³ in-plane resolution with 0.5 mm slice thickness. These orthogonal
299 2D acquisitions were reconstructed into high-resolution 3D volumes using a B-spline fitting
300 algorithm⁸⁷. Using this synthesized dataset and the CAMRI images, we created two ANTsX-
301 based population templates⁵⁸, each paired with a manually delineated brain mask. These
302 served as the basis for training an initial template-based brain extraction model. Deep
303 learning training of the network employed aggressive data augmentation strategies, including
304 bias field simulation, histogram warping, random spatial deformation, noise injection, and
305 anisotropic resampling. This enabled the model to generalize beyond the two templates.
306 The initial model was released through ANTsXNet and made publicly available.

307 Subsequent community use led to further improvements. A research group applying the
308 tool to their own ex vivo T2-weighted mouse brain data contributed a third template and
309 associated mask (acquired at 0.08 mm isotropic resolution). Incorporating this into the
310 training data improved robustness and accuracy to an independent dataset and extended
311 the model's generalizability. The refined model is distributed through ANTsPyNet via
312 `antspynet.mouse_brain_extraction(...)`.

313 **2.3.2 Template-based mouse brain anatomical labeling**

314 The AllenCCFv3 atlas and its hierarchical ontology, along with the DevCCF, provide a strong
315 foundation for developing region-wise anatomical labeling models for multi-modal mouse
316 brain imaging. Using the `allensdk` Python library, we generated a coarse segmentation
317 scheme by grouping anatomical labels into six major regions: cerebral cortex, cerebral nuclei,
318 brainstem, cerebellum, main olfactory bulb, and hippocampal formation. These labels were

mapped onto the P56 T2-weighted DevCCF template to serve as training targets. We trained a 3D U-net-based segmentation network using this template and the same augmentation strategies described for brain extraction. The model is publicly available via ANTsXNet (`antspynet.mouse_brain_parcellation(...)`) and supports robust anatomical labeling across diverse imaging geometries and contrasts. The inclusion of aggressive augmentation, including simulated anisotropy, enables the model to perform well even on thick-slice input data. Internally, the model reconstructs isotropic probability and label maps, facilitating downstream morphometric analyses. For example, this network integrates with the ANTsX cortical thickness estimation pipeline (`antspynet.mouse_cortical_thickness(...)`) to produce voxelwise cortical thickness maps, even when applied to anisotropic or limited-resolution mouse brain data.

2.3.3 Evaluation

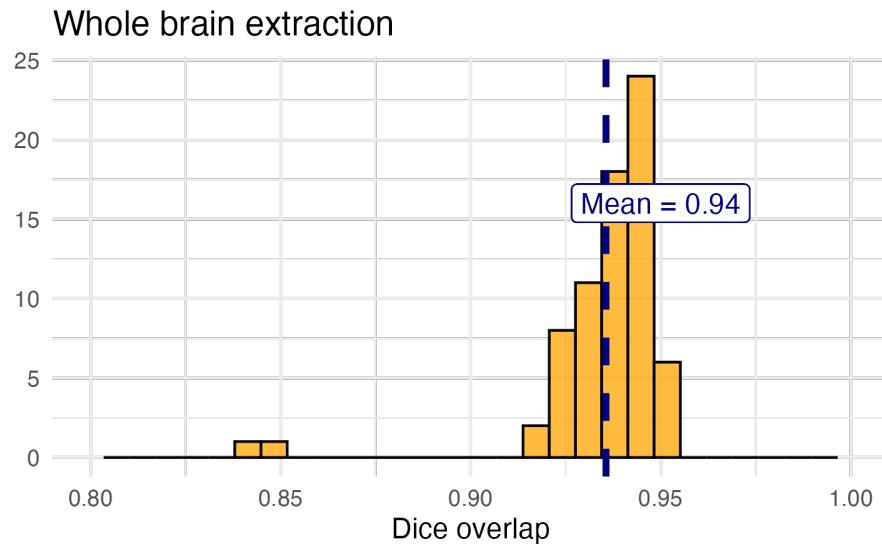
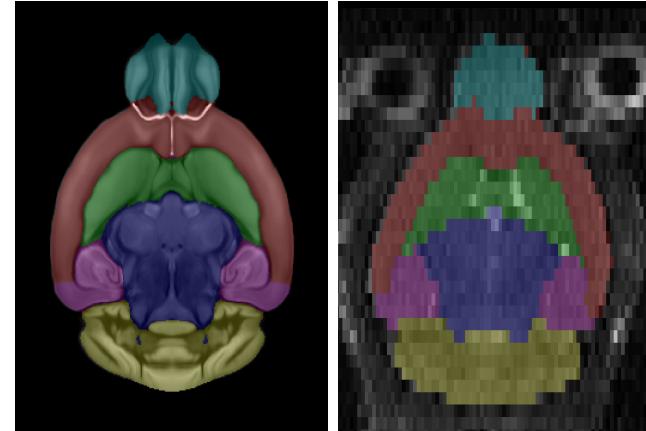


Figure 8: Evaluation of the ANTsX mouse brain extraction on an independent, publicly available dataset consisting of 12 specimens \times 7 time points = 84 total images. Dice overlap comparisons with the user-generated brain masks provide good agreement with the automated results from the brain extraction network.

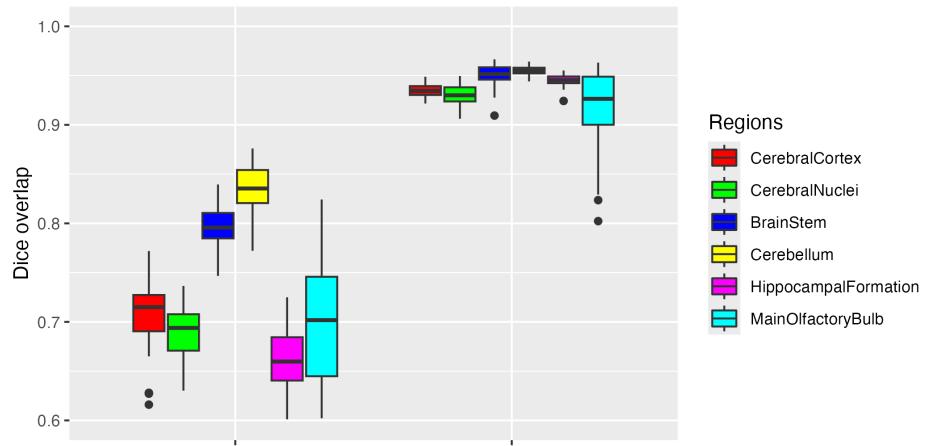
For evaluation, we used an additional publicly available dataset⁶⁹ that is completely independent from the data used in training the brain extraction and parcellation networks. Data includes 12 specimens each imaged at seven time points (Day 0, Day 3, Week 1, Week 4,



(a)

(b)

Normalization to AllenCCFv3



(c)

Figure 9: Evaluation of the ANTsX deep learning-based mouse brain parcellation on a diverse MRI cohort. (a) T2-weighted DevCCF P56 template with the six-region parcellation: cerebral cortex, nuclei, brain stem, cerebellum, main olfactory bulb, and hippocampal formation. (b) Example segmentation result from a representative subject (NR5, Day 0) using the proposed deep learning pipeline. (c) Dice overlap scores across the full evaluation cohort ($n = 84$), comparing anatomical alignment achieved via registration using intensity alone versus registration guided by the predicted parcellation. Dice values were computed using manually segmented labels transformed to AllenCCFv3 space.

³³⁴ Week 8, Week 20) with in-house-generated brain masks (i.e., produced by the data providers)
³³⁵ for a total of 84 images. Spacing is anistropic with an in-plane resolution of $0.1 \times 0.1 \text{ mm}^2$
³³⁶ and a slice thickness of 0.5 mm.

³³⁷ Figure 8 summarizes the whole-brain overlap between manually segmented reference masks
³³⁸ and the predicted segmentations for all 84 images in the evaluation cohort. The proposed
³³⁹ network demonstrates excellent performance in brain extraction across a wide age range. To
³⁴⁰ further assess the utility of the parcellation network, we used the predicted labels to guide
³⁴¹ anatomically informed registration to the AllenCCFv3 atlas using ANTsX multi-component
³⁴² registration, and compared this to intensity-only registration (Figure 9). While intensity-
³⁴³ based alignment performs reasonably well, incorporating the predicted parcellation signifi-
³⁴⁴cantly improves regional correspondence. Dice scores shown in Figure 9(c) were computed
³⁴⁵ using manually segmented labels transformed to AllenCCFv3 space.

³⁴⁶ **3 Discussion**

³⁴⁷ The diverse mouse brain cell type profiles gathered through BICCN and associated efforts
³⁴⁸ provide a rich multi-modal resource to the research community. However, despite significant
³⁴⁹ progress, optimal leveraging of these valuable resources remains an ongoing challenge. A
³⁵⁰ central component to data integration is accurately mapping novel cell type data into com-
³⁵¹ mon coordinate frameworks (CCFs) for subsequent processing and analysis. To meet these
³⁵² needs, tools for mapping mouse brain data must be both broadly accessible and capable of
³⁵³ addressing challenges unique to each modality. In this work, we described modular ANTsX-
³⁵⁴ based pipelines developed to support three distinct BICCN efforts encompassing spatial
³⁵⁵ transcriptomic, morphological, and developmental data. We demonstrated how a flexible
³⁵⁶ image analysis toolkit like ANTsX can be tailored to address specific modality-driven con-
³⁵⁷ straints by leveraging reusable, validated components.

³⁵⁸ The MERFISH mapping pipeline illustrates how ANTsX tools can be adapted to accom-
³⁵⁹ modate high-resolution spatial transcriptomic data. While the general mapping strategy is
³⁶⁰ applicable to other sectioned histological data, the pipeline includes specific adjustments for
³⁶¹ known anatomical and imaging artifacts present in MERFISH datasets. As such, this exam-
³⁶² ple demonstrates how general-purpose tools can be customized to meet the requirements of
³⁶³ highly specialized data types.

³⁶⁴ The fMOST mapping pipeline was developed with the intention of broader applicability.
³⁶⁵ Built primarily from existing ANTsX preprocessing and registration modules, this pipeline
³⁶⁶ introduces an fMOST-specific intermediate atlas to facilitate consistent mappings to the
³⁶⁷ AllenCCFv3. The use of a canonical fMOST atlas reduces the need for repeated manual
³⁶⁸ alignment across new datasets, and the resulting transformations can be directly applied to
³⁶⁹ associated single-neuron reconstructions. This supports integrative morphological analysis
³⁷⁰ across specimens using a common coordinate system.

³⁷¹ For developmental data, we introduced a velocity field-based model for continuous interpo-
³⁷² lation between discrete DevCCF timepoints. Although the DevCCF substantially expands
³⁷³ coverage of developmental stages relative to prior atlases, temporal gaps remain. The ve-

374 locity model enables spatio-temporal transformations within the full developmental interval
375 and supports the generation of virtual templates at unsampled ages. This functionality is
376 built using ANTsX components for velocity field optimization and integration, and offers
377 a novel mechanism for interpolating across the non-linear developmental trajectory of the
378 mouse brain. Such interpolation has potential utility for both anatomical harmonization and
379 longitudinal analyses. Interestingly, long-range transformations (e.g., P56 to E11.5) revealed
380 anatomy evolving in plausible ways yet sometimes diverging from known developmental pat-
381 terns (e.g., hippocampal shape changes) reflecting the input data and offering insight into
382 temporal gaps. These behaviors could assist future efforts to determine which additional
383 time points would most improve spatiotemporal coverage.

384 We also introduced a template-based deep learning pipeline for mouse brain extraction and
385 parcellation using aggressive data augmentation. This approach is designed to reduce the
386 reliance on large annotated training datasets, which remain limited in the mouse imaging
387 domain. Evaluation on independent data demonstrates promising generalization, though
388 further refinement will be necessary. As with our human-based ANTsX pipelines, failure
389 cases can be manually corrected and recycled into future training cycles. Community con-
390 tributions are welcomed and encouraged, providing a pathway for continuous improvement
391 and adaptation to new datasets.

392 The ANTsX ecosystem offers a powerful foundation for constructing scalable, reproducible
393 pipelines for mouse brain data analysis. Its modular design and multi-platform support
394 enable researchers to develop customized workflows without extensive new software devel-
395 opment. The widespread use of ANTsX components across the neuroimaging community
396 attests to its utility and reliability. As a continuation of the BICCN program, ANTsX is
397 well positioned to support the goals of the BRAIN Initiative Cell Atlas Network (BICAN)
398 and future efforts to extend these mapping strategies to the human brain.

³⁹⁹ **4 Methods**

⁴⁰⁰ The following methods are all available as part of the ANTsX ecosystem with analogous
⁴⁰¹ elements existing in both ANTsR (ANTs in R) and ANTsPy (ANTs in Python), under-
⁴⁰² pinned by a shared ANTs/ITK C++ core. Most development for the work described was
⁴⁰³ performed using ANTsPy. For equivalent functionality in ANTsR, we refer the reader to the
⁴⁰⁴ comprehensive ANTsX tutorial: <https://tinyurl.com/antsxtutorial>.

⁴⁰⁵ **4.1 General ANTsX utilities**

⁴⁰⁶ Although focused on distinct data types, the three pipelines presented in this work share
⁴⁰⁷ common components that address general challenges in mapping mouse brain data. These
⁴⁰⁸ include correcting image intensity artifacts, denoising, spatial registration, template gen-
⁴⁰⁹ eration, and visualization. Table 1 provides a concise summary of the relevant ANTsX
⁴¹⁰ functionality.

⁴¹¹ **Preprocessing: bias field correction and denoising.** Standard preprocessing steps in
⁴¹² mouse brain imaging include correcting for spatial intensity inhomogeneities and reducing im-
⁴¹³ age noise, both of which can impact registration accuracy and downstream analysis. ANTsX
⁴¹⁴ provides implementations of widely used methods for these tasks. The N4 bias field correction
⁴¹⁵ algorithm⁵¹, originally developed in ANTs and contributed to ITK, mitigates artifactual, low-
⁴¹⁶ frequency intensity variation and is accessible via `ants.n4_bias_field_correction(...)`.
⁴¹⁷ Patch-based denoising⁶⁰ has been implemented as `ants.denoise_image(...)`.

⁴¹⁸ **Image registration.** ANTsX includes a robust and flexible framework for pairwise
⁴¹⁹ and groupwise image registration⁸¹. At its core is the SyN algorithm⁵⁰, a symmetric
⁴²⁰ diffeomorphic model with optional B-spline regularization⁶⁶. In ANTsPy, registration
⁴²¹ is performed via `ants.registration(...)` using preconfigured parameter sets (e.g.,
⁴²² `antsRegistrationSyNQuick[s]`, `antsRegistrationSyN[s]`) suitable for different imaging
⁴²³ modalities and levels of computational demand. Resulting transformations can be applied
⁴²⁴ to new images with `ants.apply_transforms(...)`.

Table 1: Sampling of ANTsX functionality

<i>ANTsPy: Preprocessing</i>	
bias field correction	<code>n4_bias_field_correction(...)</code>
image denoising	<code>denoise_image(...)</code>
<i>ANTsPy: Registration</i>	
image registration	<code>registration(...)</code>
image transformation	<code>apply_transforms(...)</code>
template generation	<code>build_template(...)</code>
landmark registration	<code>fit_transform_to_paired_points(...)</code>
time-varying landmark reg.	<code>fit_time_varying_transform_to_point_sets(...)</code>
integrate velocity field	<code>integrate_velocity_field(...)</code>
invert displacement field	<code>invert_displacement_field(...)</code>
<i>ANTsPy: Segmentation</i>	
MRF-based segmentation	<code>atropos(...)</code>
Joint label fusion	<code>joint_label_fusion(...)</code>
diffeomorphic thickness	<code>kelly_kapowski(...)</code>
<i>ANTsPy: Miscellaneous</i>	
Regional intensity statistics	<code>label_stats(...)</code>
Regional shape measures	<code>label_geometry_measures(...)</code>
B-spline approximation	<code>fit_bspline_object_to_scattered_data(...)</code>
Visualize images and overlays	<code>plot(...)</code>
<i>ANTsPyNet: Mouse-specific</i>	
brain extraction	<code>mouse_brain_extraction(...modality="t2"...)</code>
brain parcellation	<code>mouse_brain_parcellation(...)</code>
cortical thickness	<code>mouse_cortical_thickness(...)</code>
super resolution	<code>mouse_histology_super_resolution(...)</code>

ANTsX provides state-of-the-art functionality for processing biomedical image data. Such tools, including deep learning networks, support a variety of mapping-related tasks. A more comprehensive listing of ANTsX tools with self-contained R and Python examples is provided as a gist page on GitHub (<https://tinyurl.com/antsxtutorial>).

425 **Template generation.** ANTsX supports population-based template generation through it-
426 erative pairwise registration to an evolving estimate of the mean shape and intensity reference
427 space across subjects⁵⁸. This functionality was used in generating the DevCCF templates¹⁶.
428 The procedure, implemented as `ants.build_template(...)`, produces average images in
429 both shape and intensity by aligning all inputs to a common evolving template.

430 **Visualization.** To support visual inspection and quality control, ANTsPy provides flexible
431 image visualization with `ants.plot(...)`. This function enables multi-slice and multi-
432 orientation rendering with optional overlays and label maps.

433 4.2 Mapping fMOST data to AllenCCFv3

434 **Preprocessing.** Mapping fMOST data into the AllenCCFv3 presents unique challenges due
435 to its native ultra-high resolution and imaging artifacts common to the fMOST modality.
436 Each fMOST image can exceed a terabyte in size, with spatial resolutions far exceeding
437 those of the AllenCCFv3 ($25\text{ }\mu\text{m}$ isotropic). To reduce computational burden and prevent
438 resolution mismatch, each fMOST image is downsampled using cubic B-spline interpolation
439 via `ants.resample_image(...)` to match the template resolution.

440 Stripe artifacts (i.e., periodic intensity distortions caused by nonuniform sectioning or il-
441 lumination) are common in fMOST and can mislead deformable registration algorithms.
442 These were removed using a custom 3D notch filter (`remove_stripe_artifact(...)`) im-
443 plemented in the `CCFAlignmentToolkit` using SciPy frequency domain filtering. The filter
444 targets dominant stripe frequencies along a user-specified axis in the Fourier domain. In
445 addition, intensity inhomogeneity across sections, often arising from variable staining or
446 illumination, was corrected using N4 bias field correction.

447 **Template-based spatial normalization.** To facilitate reproducible mapping, we first
448 constructed a contralaterally symmetric average template from 30 fMOST brains and their
449 mirrored counterparts using ANTsX template-building tools. Because the AllenCCFv3 and
450 fMOST data differ substantially in both intensity contrast and morphology, direct deformable
451 registration between individual fMOST brains and the AllenCCFv3 was insufficiently robust.

452 Instead, we performed a one-time expert-guided label-driven registration between the aver-
453 age fMOST template and AllenCCFv3. This involved sequential alignment of seven manually
454 selected anatomical regions: 1) brain mask/ventricles, 2) caudate/putamen, 3) fimbria, 4)
455 posterior choroid plexus, 5) optic chiasm, 6) anterior choroid plexus, and 7) habenular
456 commissure which were prioritized to enable coarse-to-fine correction of shape differences.
457 Once established, this fMOST-template-to-AllenCCFv3 transform was reused for all subse-
458 quent specimens. Each new fMOST brain was then registered to the average fMOST tem-
459 plate using intensity-based registration, followed by concatenation of transforms to produce
460 the final mapping into AllenCCFv3 space.

461 **Mapping neuron projections.** A key advantage of fMOST imaging is its ability to support
462 single neuron projection reconstruction across the entire brain⁷⁷. Because these reconstruc-
463 tions are stored as 3D point sets aligned to the original fMOST volume, we applied the same
464 composite transform used for image alignment to the point data using ANTsX functional-
465 ity. This enables seamless integration of cellular morphology data into AllenCCFv3 space,
466 facilitating comparative analyses across specimens.

467 4.3 Mapping MERFISH data to AllenCCFv3

468 **Preprocessing.** MERFISH data are acquired as a series of 2D tissue sections, each com-
469 prising spatially localized gene expression measurements at subcellular resolution. To enable
470 3D mapping to the AllenCCFv3, we first constructed anatomical reference images by aggre-
471 gating the number of detected transcripts per voxel across all probes within each section.
472 These 2D projections were resampled to a resolution of $10 \mu m \times 10 \mu m$ to match the in-plane
473 resolution of the AllenCCFv3.

474 Sections were coarsely aligned using manually annotated dorsal and ventral midline points,
475 allowing initial volumetric reconstruction. However, anatomical fidelity remained limited by
476 variation in section orientation, spacing, and tissue loss. To further constrain alignment and
477 enable deformable registration, we derived region-level anatomical labels directly from the
478 gene expression data.

479 **Label creation.** To assign region labels to the MERFISH data, we use a cell type clustering
480 approach previously detailed in⁴⁹. In short, manually dissected scRNAseq data was used
481 to establish the distribution of cell types present in each of the following major regions:
482 cerebellum, CTXsp, hindbrain, HPF, hypothalamus, isocortex, LSX, midbrain, OLF, PAL,
483 sAMY, STRd, STRv, thalamus and hindbrain. Clusters in the scRNA-seq dataset were then
484 used to assign similar clusters of cell types in the MERFISH data to the regions they are
485 predominantly found in the scRNA-seq data. To account for clusters that were found at
486 low frequency in regions outside its main region we calculated for each cell its 50 nearest
487 neighbors in physical space and reassigned each cell to the region annotation dominating its
488 neighborhood.

489 **Section matching via global alignment.** A major challenge was compensating for oblique
490 cutting angles and non-uniform section thickness, which distort the anatomical shape and
491 spacing of the reconstructed volume. Rather than directly warping the MERFISH data
492 into atlas space, we globally aligned the AllenCCFv3 to the MERFISH coordinate system.
493 This was done via an affine transformation followed by resampling of AllenCCFv3 sections
494 to match the number and orientation of MERFISH sections. This approach minimizes
495 interpolation artifacts in the MERFISH data and facilitates one-to-one section matching.

496 **Landmark-driven deformable alignment.** We used a 2.5D approach for fine alignment of
497 individual sections. In each MERFISH slice, deformable registration was driven by sequential
498 alignment of anatomical landmarks between the label maps derived from MERFISH and
499 AllenCCFv3. A total of nine regions—including isocortical layers 2/3, 5, and 6, the striatum,
500 hippocampus, thalamus, and medial/lateral habenula—were registered in an empirically
501 determined order. After each round, anatomical alignment was visually assessed by an
502 expert, and the next structure was selected to maximize improvement in the remaining
503 misaligned regions.

504 The final transform for each section combined the global affine alignment and the per-
505 structure deformable registrations. These were concatenated to generate a 3D mapping from
506 the original MERFISH space to the AllenCCFv3 coordinate system. Once established, the
507 composite mapping enables direct transfer of gene-level and cell-type data from MERFISH

508 into atlas space, allowing integration with other imaging and annotation datasets.

509 4.4 DevCCF velocity flow transformation model

510 The Developmental Common Coordinate Framework (DevCCF)¹⁶ provides a discrete set of
511 age-specific templates that temporally sample the developmental trajectory. To model this
512 biological progression more continuously, we introduce a velocity flow-based paradigm for in-
513 ferring diffeomorphic transformations between developmental stages. This enables anatomi-
514 cally plausible estimation of intermediate templates or mappings at arbitrary timepoints
515 between the E11.5 and P56 endpoints of the DevCCF. Our approach builds on established
516 insights from time-varying diffeomorphic registration⁶⁵, where a velocity field governs the
517 smooth deformation of anatomical structures over time. Importantly, the framework is ex-
518 tensible and can naturally accommodate additional timepoints for the potential expansion
519 of the DevCCF.

520 **Point sampling and region correspondence.** We first coalesced the anatomical labels
521 across the seven DevCCF templates (E11.5, E13.5, E15.5, E18.5, P4, P14, P56) into 26
522 common structures that could be consistently identified across development. These include
523 major brain regions such as the cortex, cerebellum, hippocampus, midbrain, and ventricles.
524 For each successive pair of templates, we performed multi-label deformable registration us-
525 ing ANTsX to generate forward and inverse transforms between anatomical label volumes.
526 From the P56 space, we randomly sampled approximately 1e6 points within and along the
527 boundaries of each labeled region and propagated them through each pairwise mapping step
528 (e.g., P56 → P14, P14 → P4, . . . , E13.5 → E11.5). This procedure created time-indexed
529 point sets tracing the spatial evolution of each region.

530 **Velocity field fitting.** Using these point sets, we fit a continuous velocity field over develop-
531 mental time using a generalized B-spline scattered data approximation method⁸⁷. The field
532 was parameterized over a log-scaled time axis to ensure finer temporal resolution during early
533 embryonic stages, where morphological changes are most rapid. Optimization proceeded for
534 approximately 125 iterations, minimizing the average Euclidean norm between transformed
535 points at each step. Ten integration points were used to ensure numerical stability. The

536 result is a smooth, differentiable vector field that defines a diffeomorphic transform between
537 any two timepoints within the template range.

538 **Applications and availability.** This velocity model can be used to estimate spa-
539 tial transformations between any pair of developmental stages—even those for which
540 no empirical template exists—allowing researchers to create interpolated atlases, align
541 new datasets, or measure continuous structural changes. It also enables developmental
542 alignment of multi-modal data (e.g., MRI to LSFM) by acting as a unifying spatiotem-
543 poral scaffold. The underlying components for velocity field fitting and integration
544 are implemented in ITK, and the complete workflow is accessible in both ANTsPy
545 (`ants.fit_time_varying_transform_to_point_sets(...)`) and ANTsR. In addition
546 the availability of the DevCCF use case, self-contained examples and usage tutorials are
547 provided in our public codebase.

548 4.5 Automated brain extraction and parcellation with ANTsXNet

549 To support template-based deep learning approaches for structural brain extraction and par-
550 cellation, we implemented dedicated pipelines using the ANTsXNet framework. ANTsXNet
551 comprises open-source deep learning libraries in both Python (ANTsPyNet) and R (ANTsR-
552 Net) that interface with the broader ANTsX ecosystem and are built on TensorFlow/Keras.
553 Our mouse brain pipelines mirror existing ANTsXNet tools for human imaging but are
554 adapted for species-specific anatomical variation, lower SNR, and heterogeneous acquisition
555 protocols.

556 4.5.1 Deep learning training setup

557 All network-based approaches were implemented using a standard U-net⁸⁸ architecture and
558 hyperparameters previously evaluated in ANTsXNet pipelines for human brain imaging⁴⁵.
559 This design follows the ‘no-new-net’ principle⁸⁹, which demonstrates that a well-configured,
560 conventional U-net can achieve robust and competitive performance across a wide range of
561 biomedical segmentation tasks with little to no architectural modifications from the original.

562 Both networks use a 3D U-net architecture implemented in TensorFlow/Keras, with five
563 encoding/decoding levels and skip connections. The loss function combined Dice and cate-
564 gorical cross-entropy terms. Training used a batch size of 4, Adam optimizer with an initial
565 learning rate of 2e-4, and early stopping based on validation loss. Training was performed on
566 an NVIDIA DGX system ($4 \times$ Tesla V100 GPUs, 256 GB RAM). Model weights and prepro-
567 cessing routines are shared across ANTsPyNet and ANTsRNet to ensure reproducibility and
568 language portability. For both published and unpublished trained networks available through
569 ANTsXNet, all training scripts and data augmentation generators are publicly available at
570 <https://github.com/ntustison/ANTsXNetTraining>.

571 **Data augmentation.** Robust data augmentation was critical to generalization across scan-
572 ners, contrast types, and resolutions. We applied both intensity- and shape-based augmen-
573 tation strategies:

574 • *Intensity augmentations:*

- 575 – Gaussian, Poisson, and salt-and-pepper noise:
576 `ants.add_noise_to_image(...)`
577 – Simulated intensity inhomogeneity via bias field modeling⁵¹:
578 `antspynet.simulate_bias_field(...)`
579 – Histogram warping to simulate contrast variation⁹⁰:
580 `antspynet.histogram_warp_image_intensities(...)`

581 • *Shape augmentations:*

- 582 – Random nonlinear deformations and affine transforms:
583 `antspynet.randomly_transform_image_data(...)`
584 – Anisotropic resampling across axial, sagittal, and coronal planes:
585 `ants.resample_image(...)`

586 **4.5.2 Brain extraction**

587 We originally trained a mouse-specific brain extraction model on two manually masked
588 T2-weighted templates, generated from public datasets^{67,68}. One of the templates was
589 constructed from orthogonal 2D acquisitions using B-spline-based volumetric synthesis via
590 `ants.fit_bspline_object_to_scattered_data(...)`. Normalized gradient magnitude
591 was used as a weighting function to emphasize boundaries during reconstruction⁸⁷.

592 This training strategy provides strong spatial priors despite limited data by leveraging high-
593 quality template images and aggressive augmentation to mimic population variability. Dur-
594 ing the development of this work, the network was further refined through community en-
595 gagement. A user from a U.S.-based research institute applied this publicly available (but
596 then unpublished) brain extraction tool to their own mouse MRI dataset. Based on feedback
597 and iterative collaboration with the ANTsX team, the model was retrained and improved to
598 better generalize to additional imaging contexts. This reflects our broader commitment to
599 community-driven development and responsiveness to user needs across diverse mouse brain
600 imaging scenarios.

601 The final trained network is available via ANTsXNet through the function
602 `antspynet.mouse_extraction(...)`. Additionally, both template/mask pairs are
603 accessible via ANTsXNet. For example, one such image pair is available via:

- 604 • Template:

```
605     antspynet.get_antsxnet_data("bsplineT2MouseTemplate")
```

- 606 • Brain mask:

```
607     antspynet.get_antsxnet_data("bsplineT2MouseTemplateBrainMask")
```

608 **4.5.3 Brain parcellation**

609 For brain parcellation, we trained a 3D U-net model using the DevCCF P56 T2-weighted
610 template and anatomical segmentations derived from AllenCCFv3. This template-based
611 training strategy enables the model to produce accurate, multi-region parcellations without
612 requiring large-scale annotated subject data.

613 To normalize intensity across specimens, input images were preprocessed using rank-based
614 intensity normalization (`ants.rank_intensity(...)`). Spatial harmonization was achieved
615 through affine and deformable alignment of each extracted brain to the P56 template prior
616 to inference. In addition to the normalized image input, the network also receives prior
617 probability maps derived from the atlas segmentations, providing additional spatial context.
618 This general parcellation deep learning framework has also been applied in collaboration
619 with other groups pursuing related but distinct projects. In one case, a model variant was
620 adapted for T2-weighted MRI using an alternative anatomical labeling scheme; in another,
621 a separate model was developed for serial two-photon tomography (STPT) with a different
622 parcellation set. All three models are accessible through a shared interface in ANTsXNet:
623 `antspynet.mouse_brain_parcellation(...)`. Ongoing work is further extending this ap-
624 proach to embryonic mouse brain data. These independent efforts reflect broader community
625 interest in adaptable parcellation tools and reinforce the utility of ANTsXNet as a platform
626 for reproducible, extensible deep learning workflows.

627 4.5.4 Evaluation and reuse

628 To assess model generalizability, both the brain extraction and parcellation networks were
629 evaluated on an independent longitudinal dataset comprising multiple imaging sessions with
630 varied acquisition parameters⁶⁹. Although each label or imaging modality required re-
631 training, the process was streamlined by the reusable ANTsX infrastructure enabled by
632 rapid adaptation with minimal overhead. These results illustrate the practical benefits of a
633 template-based, low-shot strategy and modular deep learning framework. All trained mod-
634 els, associated training scripts, and supporting resources are openly available and designed
635 for straightforward integration into ANTsX workflows.

636 **Data availability**

637 All data and software used in this work are publicly available. The DevCCF atlas is
638 available at <https://kimlab.io/brain-map/DevCCF/>. ANTsPy, ANTsR, ANTsPyNet, and
639 ANTsRNet are available through GitHub at the ANTsX Ecosystem ([https://github.com/](https://github.com/ANTsX)
640 [ANTsX](#)). Training scripts for all deep learning functionality in ANTsXNet can also be
641 found on GitHub (<https://github.com/ntustison/ANTsXNetTraining>). A GitHub reposi-
642 tory specifically pertaining to the AllenCCFv3 mapping is available at <https://github.com/>
643 [dontminchenit/CCFAAlignmentToolkit](#). For the other two contributions contained in this
644 work, the longitudinal DevCCF mapping and mouse cortical thickness pipeline, we refer the
645 interested reader to <https://github.com/ntustison/ANTsXMouseBrainMapping>.

646 Acknowledgments

647 Support for the research reported in this work includes funding from the National Institute
648 of Biomedical Imaging and Bioengineering (R01-EB031722) and National Institute of Mental
649 Health (RF1-MH124605 and U24-MH114827).

650 We also acknowledge the data contribution of Dr. Adam Raikes (GitHub @araikes) of the
651 Center for Innovation in Brain Science at the University of Arizona for refining the weights
652 of the mouse brain extraction network.

653 **Author contributions**

654 N.T., M.C., and J.G. wrote the main manuscript text and figures. M.C., M.K., R.D., S.S.,
655 Q.W., L.G., J.D., C.G., and J.G. developed the Allen registration pipelines. N.T. and F.K.
656 developed the time-varying velocity transformation model for the DevCCF. N.T. and M.T.
657 developed the brain parcellation and cortical thickness methodology. All authors reviewed
658 the manuscript.

659 **References**

- 660 1. Keller, P. J. & Ahrens, M. B. Visualizing whole-brain activity and development at
the single-cell level using light-sheet microscopy. *Neuron* **85**, 462–83 (2015).
- 661 2. La Manno, G. *et al.* Molecular architecture of the developing mouse brain. *Nature*
596, 92–96 (2021).
- 662 3. Wen, L. *et al.* Single-cell technologies: From research to application. *Innovation
(Camb)* **3**, 100342 (2022).
- 663 4. Oh, S. W. *et al.* A mesoscale connectome of the mouse brain. *Nature* **508**, 207–14
(2014).
- 664 5. Gong, H. *et al.* Continuously tracing brain-wide long-distance axonal projections in
mice at a one-micron voxel resolution. *Neuroimage* **74**, 87–98 (2013).
- 665 6. Li, A. *et al.* Micro-optical sectioning tomography to obtain a high-resolution atlas of
the mouse brain. *Science* **330**, 1404–8 (2010).
- 666 7. Ueda, H. R. *et al.* Tissue clearing and its applications in neuroscience. *Nat Rev
Neurosci* **21**, 61–79 (2020).
- 667 8. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by
spatial transcriptomics. *Science* **353**, 78–82 (2016).
- 668 9. Burgess, D. J. Spatial transcriptomics coming of age. *Nat Rev Genet* **20**, 317 (2019).
- 669 10. Hardwick, S. A. *et al.* Single-nuclei isoform RNA sequencing unlocks barcoded exon
connectivity in frozen brain tissue. *Nature biotechnology* **40**, 1082–1092 (2022).
- 670 11. Hawrylycz, M. *et al.* A guide to the BRAIN initiative cell census network data
ecosystem. *PLoS biology* **21**, e3002133 (2023).
- 671 12. Wang, Q. *et al.* The allen mouse brain common coordinate framework: A 3D reference
atlas. *Cell* **181**, 936–953.e20 (2020).
- 672 13. Perens, J. *et al.* An optimized mouse brain atlas for automated mapping and quantification
of neuronal activity using iDISCO+ and light sheet fluorescence microscopy.
Neuroinformatics **19**, 433–446 (2021).
- 673 14. Ma, Y. *et al.* A three-dimensional digital atlas database of the adult C57BL/6J mouse
brain by magnetic resonance microscopy. *Neuroscience* **135**, 1203–1215 (2005).

- 674 15. Qu, L. *et al.* Cross-modal coherent registration of whole mouse brains. *Nature Methods* **19**, 111–118 (2022).
- 675 16. Kronman, F. N. *et al.* [Developmental mouse brain common coordinate framework](#). *Nat Commun* **15**, 9072 (2024).
- 676 17. Chuang, N. *et al.* An MRI-based atlas and database of the developing mouse brain. *Neuroimage* **54**, 80–89 (2011).
- 677 18. Dries, R. *et al.* Advances in spatial transcriptomic data analysis. *Genome research* **31**, 1706–1718 (2021).
- 678 19. Ricci, P. *et al.* Removing striping artifacts in light-sheet fluorescence microscopy: A review. *Progress in biophysics and molecular biology* **168**, 52–65 (2022).
- 679 20. Agarwal, N., Xu, X. & Gopi, M. Robust registration of mouse brain slices with severe histological artifacts. in *Proceedings of the tenth indian conference on computer vision, graphics and image processing* 1–8 (2016).
- 680 21. Agarwal, N., Xu, X. & Gopi, M. Automatic detection of histological artifacts in mouse brain slice images. in *Medical computer vision and bayesian and graphical models for biomedical imaging: MICCAI 2016 international workshops, MCV and BAMBI, athens, greece, october 21, 2016, revised selected papers* 8 105–115 (Springer, 2017).
- 681 22. Tward, D. *et al.* 3d mapping of serial histology sections with anomalies using a novel robust deformable registration algorithm. in *International workshop on multimodal brain image analysis* 162–173 (Springer, 2019).
- 682 23. Cahill, L. S. *et al.* Preparation of fixed mouse brains for MRI. *Neuroimage* **60**, 933–939 (2012).
- 683 24. Biancalani, T. *et al.* [Deep learning and alignment of spatially resolved single-cell transcriptomes with tangram](#). *Nat Methods* **18**, 1352–1362 (2021).
- 684 25. Sunkin, S. M. *et al.* Allen brain atlas: An integrated spatio-temporal portal for exploring the central nervous system. *Nucleic acids research* **41**, D996–D1008 (2012).
- 685 26. Kim, Y. *et al.* Brain-wide maps reveal stereotyped cell-type-based cortical architecture and subcortical sexual dimorphism. *Cell* **171**, 456–469 (2017).

- 686 27. Fürth, D. *et al.* An interactive framework for whole-brain maps at cellular resolution. *Nat Neurosci* **21**, 139–149 (2018).
- 687 28. Li, Y. *et al.* mBrainAligner-web: A web server for cross-modal coherent registration of whole mouse brains. *Bioinformatics* **38**, 4654–4655 (2022).
- 688 29. Puchades, M. A., Csucs, G., Ledergerber, D., Leergaard, T. B. & Bjaalie, J. G. Spatial registration of serial microscopic brain images to three-dimensional reference atlases with the QuickNII tool. *PloS one* **14**, e0216796 (2019).
- 689 30. Eastwood, B. S. *et al.* Whole mouse brain reconstruction and registration to a reference atlas with standard histochemical processing of coronal sections. *Journal of Comparative Neurology* **527**, 2170–2178 (2019).
- 690 31. Ni, H. *et al.* A robust image registration interface for large volume brain atlas. *Sci Rep* **10**, 2139 (2020).
- 691 32. Pallast, N. *et al.* Processing pipeline for atlas-based imaging data analysis of structural and functional mouse brain MRI (AIDAmri). *Front Neuroinform* **13**, 42 (2019).
- 692 33. Celestine, M., Nadkarni, N. A., Garin, C. M., Bougacha, S. & Dhenain, M. Sammba-MRI: A library for processing SmAll-MaMmal BrAin MRI data in python. *Front Neuroinform* **14**, 24 (2020).
- 693 34. Ioanas, H.-I., Marks, M., Zerbi, V., Yanik, M. F. & Rudin, M. An optimized registration workflow and standard geometric space for small animal brain imaging. *Neuroimage* **241**, 118386 (2021).
- 694 35. Aggarwal, M., Zhang, J., Miller, M. I., Sidman, R. L. & Mori, S. Magnetic resonance imaging and micro-computed tomography combined atlas of developing and adult mouse brains for stereotaxic surgery. *Neuroscience* **162**, 1339–1350 (2009).
- 695 36. Chandrashekhar, V. *et al.* CloudReg: Automatic terabyte-scale cross-modal brain volume registration. *Nature methods* **18**, 845–846 (2021).
- 696 37. Jin, M. *et al.* SMART: An open-source extension of WholeBrain for intact mouse brain registration and segmentation. *eNeuro* **9**, (2022).
- 697 38. Negwer, M. *et al.* FriendlyClearMap: An optimized toolkit for mouse brain mapping and analysis. *Gigascience* **12**, (2022).

- 698 39. Lin, W. *et al.* Whole-brain mapping of histaminergic projections in mouse brain. *Proceedings of the National Academy of Sciences* **120**, e2216231120 (2023).
- 699 40. Zhang, M. *et al.* Spatially resolved cell atlas of the mouse primary motor cortex by MERFISH. *Nature* **598**, 137–143 (2021).
- 700 41. Shi, H. *et al.* Spatial atlas of the mouse central nervous system at molecular resolution. *Nature* **622**, 552–561 (2023).
- 701 42. Zhang, Y. *et al.* Reference-based cell type matching of *in situ* image-based spatial transcriptomics data on primary visual cortex of mouse brain. *Scientific Reports* **13**, 9567 (2023).
- 702 43. Klein, S., Staring, M., Murphy, K., Viergever, M. A. & Pluim, J. P. W. [Elastix: A toolbox for intensity-based medical image registration](#). *IEEE Trans Med Imaging* **29**, 196–205 (2010).
- 703 44. Fedorov, A. *et al.* 3D slicer as an image computing platform for the quantitative imaging network. *Magnetic resonance imaging* **30**, 1323–1341 (2012).
- 704 45. Tustison, N. J. *et al.* [The ANTsX ecosystem for quantitative biological and medical imaging](#). *Sci Rep* **11**, 9068 (2021).
- 705 46. Pagani, M., Damiano, M., Galbusera, A., Tsafaris, S. A. & Gozzi, A. Semi-automated registration-based anatomical labelling, voxel based morphometry and cortical thickness mapping of the mouse brain. *Journal of neuroscience methods* **267**, 62–73 (2016).
- 706 47. Anderson, R. J. *et al.* [Small animal multivariate brain analysis \(SAMBA\) - a high throughput pipeline with a validation framework](#). *Neuroinformatics* **17**, 451–472 (2019).
- 707 48. Allan Johnson, G. *et al.* Whole mouse brain connectomics. *Journal of Comparative Neurology* **527**, 2146–2157 (2019).
- 708 49. Yao, Z. *et al.* [A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain](#). *Nature* **624**, 317–332 (2023).
- 709 50. Avants, B. B., Epstein, C. L., Grossman, M. & Gee, J. C. [Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain](#). *Med Image Anal* **12**, 26–41 (2008).

- 710 51. Tustison, N. J. *et al.* N4ITK: Improved N3 bias correction. *IEEE Trans Med Imaging* **29**, 1310–20 (2010).
- 711 52. Bajcsy, R. & Broit, C. Matching of deformed images. in *Sixth International Conference on Pattern Recognition (ICPR'82)* 351–353 (1982).
- 712 53. Bajcsy, R. & Kovacic, S. Multiresolution elastic matching. *Computer Vision, Graphics, and Image Processing* **46**, 1–21 (1989).
- 713 54. Gee, J. C., Reivich, M. & Bajcsy, R. Elastically deforming 3D atlas to match anatomical brain images. *J Comput Assist Tomogr* **17**, 225–36 (1993).
- 714 55. Klein, A. *et al.* Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* **46**, 786–802 (2009).
- 715 56. Murphy, K. *et al.* Evaluation of registration methods on thoracic CT: The EMPIRE10 challenge. *IEEE Trans Med Imaging* **30**, 1901–20 (2011).
- 716 57. Baheti, B. *et al.* The brain tumor sequence registration challenge: Establishing correspondence between pre-operative and follow-up MRI scans of diffuse glioma patients. (2021).
- 717 58. Avants, B. B. *et al.* The optimal template effect in hippocampus studies of diseased populations. *Neuroimage* **49**, 2457–66 (2010).
- 718 59. Avants, B. B., Tustison, N. J., Wu, J., Cook, P. A. & Gee, J. C. An open source multivariate framework for n-tissue segmentation with evaluation on public data. *Neuroinformatics* **9**, 381–400 (2011).
- 719 60. Manjón, J. V., Coupé, P., Martí-Bonmatí, L., Collins, D. L. & Robles, M. Adaptive non-local means denoising of MR images with spatially varying noise levels. *J Magn Reson Imaging* **31**, 192–203 (2010).
- 720 61. Wang, H. *et al.* Multi-atlas segmentation with joint label fusion. *IEEE Trans Pattern Anal Mach Intell* **35**, 611–23 (2013).
- 721 62. Tustison, N. J. *et al.* Optimal symmetric multimodal templates and concatenated random forests for supervised brain tumor segmentation (simplified) with ANTsR. *Neuroinformatics* (2014) doi:10.1007/s12021-014-9245-2.

- 722 63. Tustison, N. J., Yang, Y. & Salerno, M. [Advanced normalization tools for cardiac motion correction](#). in *Statistical atlases and computational models of the heart - imaging and modelling challenges* (eds. Camara, O. et al.) vol. 8896 3–12 (Springer International Publishing, 2015).
- 723 64. McCormick, M., Liu, X., Jomier, J., Marion, C. & Ibanez, L. [ITK: Enabling reproducible research and open science](#). *Front Neuroinform* **8**, 13 (2014).
- 724 65. Beg, M. F., Miller, M. I., Trouvé, A. & Younes, L. [Computing large deformation metric mappings via geodesic flows of diffeomorphisms](#). *International Journal of Computer Vision* **61**, 139–157 (2005).
- 725 66. Tustison, N. J. & Avants, B. B. [Explicit B-spline regularization in diffeomorphic image registration](#). *Front Neuroinform* **7**, 39 (2013).
- 726 67. Hsu, L.-M. *et al.* CAMRI mouse brain MRI data.
- 727 68. Reshetnikov, V. *et al.* High-resolution MRI data of brain C57BL/6 and BTBR mice in three different anatomical views.
- 728 69. Rahman, N., Xu, K., Budde, M. D., Brown, A. & Baron, C. A. [A longitudinal microstructural MRI dataset in healthy C57Bl/6 mice at 9.4 tesla](#). *Sci Data* **10**, 94 (2023).
- 729 70. Liu, J. *et al.* [Concordance of MERFISH spatial transcriptomics with bulk and single-cell RNA sequencing](#). *Life Sci Alliance* **6**, (2023).
- 730 71. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. [Cellpose: A generalist algorithm for cellular segmentation](#). *Nat Methods* **18**, 100–106 (2021).
- 731 72. Jia, H., Yap, P.-T., Wu, G., Wang, Q. & Shen, D. Intermediate templates guided groupwise registration of diffusion tensor images. *NeuroImage* **54**, 928–939 (2011).
- 732 73. Tang, S., Fan, Y., Wu, G., Kim, M. & Shen, D. RABBIT: Rapid alignment of brains by building intermediate templates. *NeuroImage* **47**, 1277–1287 (2009).
- 733 74. Dewey, B. E., Carass, A., Blitz, A. M. & Prince, J. L. Efficient multi-atlas registration using an intermediate template image. in *Proceedings of SPIE—the international society for optical engineering* vol. 10137 (NIH Public Access, 2017).

- 734 75. Perens, J. *et al.* Multimodal 3D mouse brain atlas framework with the skull-derived
coordinate system. *Neuroinformatics* **21**, 269–286 (2023).
- 735 76. Rotolo, T., Smallwood, P. M., Williams, J. & Nathans, J. Genetically-directed, cell
type-specific sparse labeling for the analysis of neuronal morphology. *PLoS One* **3**,
e4099 (2008).
- 736 77. Peng, H. *et al.* Morphological diversity of single neurons in molecularly defined cell
types. *Nature* **598**, 174–181 (2021).
- 737 78. Gong, H. *et al.* High-throughput dual-colour precision imaging for brain-wide connec-
tome with cytoarchitectonic landmarks at the cellular level. *Nat Commun* **7**, 12142
(2016).
- 738 79. Wang, J. *et al.* Divergent projection patterns revealed by reconstruction of individual
neurons in orbitofrontal cortex. *Neurosci Bull* **37**, 461–477 (2021).
- 739 80. Kronman, F. A. *et al.* Developmental mouse brain common coordinate framework.
bioRxiv (2023) doi:[10.1101/2023.09.14.557789](https://doi.org/10.1101/2023.09.14.557789).
- 740 81. Avants, B. B. *et al.* The Insight ToolKit image registration framework. *Front Neu-
roinform* **8**, 44 (2014).
- 741 82. Chon, U., Vanselow, D. J., Cheng, K. C. & Kim, Y. Enhanced and unified anatomical
labeling for a common mouse brain atlas. *Nat Commun* **10**, 5067 (2019).
- 742 83. Tasic, B. *et al.* Adult mouse cortical cell taxonomy revealed by single cell transcrip-
omics. *Nat Neurosci* **19**, 335–46 (2016).
- 743 84. Bergmann, E., Gofman, X., Kavushansky, A. & Kahn, I. Individual variability in
functional connectivity architecture of the mouse brain. *Commun Biol* **3**, 738 (2020).
- 744 85. Billot, B. *et al.* SynthSeg: Segmentation of brain MRI scans of any contrast and
resolution without retraining. *Med Image Anal* **86**, 102789 (2023).
- 745 86. Rolfe, S. M., Whikehart, S. M. & Maga, A. M. Deep learning enabled multi-organ
segmentation of mouse embryos. *Biol Open* **12**, bio059698 (2023).
- 746 87. Tustison, N. J. & Amini, A. A. Biventricular myocardial strains via nonrigid regis-
tration of anatomical NURBS model [corrected]. *IEEE Trans Med Imaging* **25**, 94–112
(2006).

- 747 88. Falk, T. *et al.* U-net: Deep learning for cell counting, detection, and morphometry. *Nat Methods* **16**, 67–70 (2019).
- 748 89. Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J. & Maier-Hein, K. H. nnU-net: A self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods* **18**, 203–211 (2021).
- 749 90. Tustison, N. J. *et al.* Image- versus histogram-based considerations in semantic segmentation of pulmonary hyperpolarized gas images. *Magn Reson Med* **86**, 2822–2836 (2021).