An active texture-based digital atlas enables automated mapping of structures and markers across brains

Yuncong Chen[1], **Lauren E. McElvain**[2], **Alexander S. Tolpygo**[3], **Daniel Ferrante**[3], **Beth Friedman**[1], **Partha P. Mitra**[3], **Harvey J. Karten**[4], **Yoav Freund**[1,6], **David Kleinfeld**[2,5,6]

[1]Department of Computer Science and Engineering, University of California, San Diego, CA

[2]Department of Physics, University of California, San Diego, CA

[3]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

[4]Department of Neurosciences, University of California School of Medicine, San Diego, CA

[5]Section of Neurobiology, University of California, San Diego, CA

Abstract

Brain atlases enable the mapping of labeled cells and newly defined projections from different brains onto a standard coordinate system. We address two fundamental issues in the construction and use of atlases. First, expert neuroanatomists ascertain the fine-scale structure of brain tissue, the "texture" formed by cell structure and organization, to define cytoarchitectural borders. Can this approach be automated, so that a machine can locate landmark structures and automatically align new brains to a reference atlas? We achieve this goal with a robust procedure that is driven by machine learning and bootstrapped from brains annotated by experts. Second, can one construct a brain atlas that is active, i.e., augmented and improved with each use? We show that the alignment of new brains to a reference atlas can continuously refine the coordinate system and associated variance. We apply this approach to the adult murine brainstem and achieve a precise alignment of projections in cytoarchitecturally ill-defined regions across brains from different animals.

Introduction

Brain atlases provide a visual depository for the everexpanding studies of neuron wiring and function [1,2]. The navigability of any atlas depends on demarcation of regional boundaries, or landmarks. The modern standard for brain atlas construction is to utilize sets of landmarks, shared across brains, to define a reference atlas [3–5] to register data from new subject brains to a common standard. The use of landmarks also provides a framework for triangulation, so that newly discovered functional brain subregions can be incorporated into the atlas [6]. Traditionally, landmark recognition has depended on skilled assessment of

brain cytoarchitecture by expert anatomists [7–9]. The primary data typically takes the form of Nissl stained histological sections that capture the texture of neural tissue [10], including such high resolution features as cell shape, size, orientation and packing density. These cytoarchitectural features have enabled discrimination of brain regions with sharp borders, such as many cranial nerve motor nuclei and cortical lamina, as well as discrimination of small nuclei with more subtle boundaries such as the nucleus ambiguus.

Landmark assignment in magnetic resonant imaging (MRI) reference brain atlases is necessarily based on low resolution images where boundaries are determined from large shifts in grey levels; more recent experimental brain atlases have adopted these standards in part to retain a compatible modality with a three-dimensional reference space dictated by magnetic resonant imaging (MRI) of a representative experimental brain [5]. This approach has long been known to necessarily limit the types of landmarks that can be used for navigation, as regions with subtle boundaries are not recognized; in part for these reasons numerous small brainstem structures in mice have not been absorbed into a standardized reference atlas. Additionally, the use of single fixed reference atlases does not incorporate the expected variance in brain regions of subject mice brains, even though it is known that brains of inbred mice can differ in the structural characteristics of neurons within a common region [11]. Thus the questions arises as to how to align structures across brains as well as how to evaluate the goodness-of-fit of the alignment across brains. In particular, the need to quantify and preserve the variation among brains calls for a probabilistic approach during the addition of new data into an updatable reference atlas. With this goal, an idealized atlas is a dynamic document that incorporates a diversity of landmark structures and also progressively improves in accuracy and resolution through the addition of new brains. This dynamic document is termed an active atlas.

Active atlases have provided a fruitful approach to collate MRI studies of high contrast brain structures in patient populations. However, the ability to chart ill-defined brain regions will demand access to the higher-resolution spatial information, such as found from optical imaging of brains [12–14]. Toward this goal, we demonstrate a software system that functions as an active atlas and is based on automated detection of brain textures. A supervised approach is adopted to create texture classifiers that will be used to identify landmarks and, further, to bootstrap a reference atlas. The texture classifiers are initialized by human expert annotators. The automated alignment process of a new brain with the reference atlas is based on machinegenerated detection of multiple landmarks in the new brain using the texture classifiers (Figure 1a–d). Final verification is performed by a human. Thus the software system serves to align new brains to a standard coordinate system that is derived from the reference atlas. The new brain then is used to update and improve the reference atlas. This process amortizes the time of expert anatomists. While experts may spend a relatively long time annotating each brain, the verification step will take only a small fraction of that time. Our end product also provides a means to use landmarks to triangulate regions with subtle ill-defined borders and then coalign such regions across separate brains with both high precision and known uncertainty (Figure 1e).

We apply our approach to the murine brainstem, i.e., the hindbrain and midbrain, across a cohort of mice. The brainstem is a challenging region to map for several reasons. Its

mechanical floppiness complicates brain positioning for imaging and sectioning. While its cytoarchitecture is marked by well -delineated cranial nerve nuclei, it is also home to premotor neurons populations in subregions with at best, subtle borders [15]. These premotor regions, mainly in the extensive reticular formation of the brainstem, are of crucial importance in regulation of brainstem output functions that range from breathing to orofacial sensorimotor behaviors [16]. Thus, issues of automation aside, the failure to form a reliable atlas of this region has stifled comparisons of studies across brains and from different laboratories. We emphasize that our approach is general. It may be applied across the entire brain. It should be useful for all brains in nervous systems that are not wholly characterized by identified enumerated neurons [17–20].

Results

Our focus is on the use of brain texture as a means to identify landmarks for the alignment of brains. Toward this goal, we used P56, male C57BL/6 mice. Fixed cryoprotected brains were sectioned in a sagittal plane on a cryostat and the quality of our histological sections was maximized through the use of an improved Cryojane™ tape-transfer method [21]. This procedure uses a supporting film during cutting and mounting to minimize physical distortion of thin sections and facilitates reliable collection of all serial sections across an entire brain of a mouse. We stained the Nissl substance, i.e., ribosomal and message RNA, which highlights neural texture across the brain.

Initialization of an active atlas

Expert anatomists were asked to bilaterally mark boundaries for a set of landmark structures (Figure 2a,b; Supplemental Figure 7). The processes is aided by a display of the annotation that resections the data in the two alternate planes in real-time (Supplemental Figure 8). The annotated data serves two purposes. One is to form a training set for our texture-based classifiers. The second is to capture the location and approximate shape of each of the landmarks and bootstrap the reference atlas. In practice, our experts contoured around each of 51 landmarks across three brains (Figure 2c,d); these correspond to 28 different structures (Table 1). Note that the right and left sides of five structures that border the midline, e.g., hypoglossal nucleus (12N), inferior (IC) and superior (SC) colliculus, area postrema (AP), and and reticulotegmental nucleus (RtTg), were fused into a single landmarks.

Training structure-specific texture classifiers—We divide the image of each brain section into overlapping square patches that are sufficiently large to contain many cells but small enough so that each landmark is tiled by many patches. For concreteness, we choose the training patches to be 100 μm in edge with a pitch of 30 μm. Patches within an annotated landmark are labeled positive, i.e., $y_m = +1$ for the m-th patch, while patches within a boundary region that surrounds the landmark are labeled negative, i.e., $y_m = -1$ (Figure 2e). The textural information of each image patch is encoded as a set of numbers, called a *feature vector* and denoted by x_m. We used a convolutional neural network (CNN) with fixed weights, i.e., the blue channel only of the Inception-BN [22] that was trained on natural scenes, to perform the encoding. The rich internal filters appear to effectively represent

histological textures in terms of a 1024-dimensional vector that defines x_m, so that each patch is represented by the pair $(\mathbf{x}_m, y_m)$.

Supervised learning is used to create the texture-based classifiers, one for each landmark and denoted $f^{landmark}$. The classifiers enable us to compute, for a given feature vector $\mathbf{x}$, the conditional probability that the corresponding patch is inside any one of the landmarks (Figure 2f). We use logistic regression, a generalized linear model, as the functional form of our classifier. The logistic function for a given landmark is defined by a weight vector, $\mathbf{w}^{landmark}$ and an offset $\theta^{landmark}$. Formally, the logistic function is used to compute the conditional probability of the label y_m for each landmark given the feature vector $\mathbf{x}_m$ for each patch, i.e.,

$$f^{landmark}(\mathbf{x}_m) = \frac{1}{1 + e^{-\left(y_m \mathbf{x}_m \cdot \mathbf{w}^{landmark} - \theta^{landmark}\right)}} \quad (1)$$

and is a number between 0 and 1. The weight vectors and offsets are found by maximizing the likelihood of the training data. The complete set of classifiers, parameterized by $\mathbf{w}^{landmark}$s and $\theta^{landmark}$, enable us to to score a new brain for the probability, $f^{landmark}(x_n)$, that the n-th patch belongs to each of the landmarks. Operationally, the classifiers represent the knowledge of experts that has been captured through machine learning, so that expertise outlives the expert.

We assessed the performance of each classifier $f^{landmark}$ in correctly predicting a landmark by a single number, the area under the receiver operator characteristic (ROC) curve. We used 1,000 positive and 1,000 negative patches from each of the annotated brains, chosen at random and split as training and testing sets. The area under the ROC curves ranged from 0.85 to 0.98 (Supplemental Figure 9) with a means of 0.92, compared to a random value of 0.50 and a maximum of 1.00.

Bootstrapping the reference atlas—The contours for each of the landmarks are interpolated to form three-dimensional volumetric annotations that jointly constitute a labeled volume for each annotated brain. The labeled volumes of all annotated brains are co-aligned and the mean and covariance of the coordinates of the centroid for each landmark are computed (Supplemental Figure 10a,b) We further derive a probabilistic volume for each landmark, denoted $p^{landmark}$, to represent the average shape by registering all three-dimensional annotations of the same landmark across all of the brains (Supplemental Figure 10c–f). We label the regions that are included in the annotations of all brains by $p^{landmark} = 1$, while regions that are incorporated by only a fraction of the annotations have $p^{landmark} < 1$. The combination of the average shapes and mean centroids of all the landmarks gives rise to the initial probabilistic reference atlas (Figure 2c,d; Supplementary Figure 10).

Automated alignment of a new brain with reference atlas

We use the trained classifiers and the reference atlas (Figure 2c,d) to align a new serially sectioned brain with the reference atlas. We consider first the use of additional thionin

counterstained sections to test the accuracy and reproducibility of our approach (Figures 1a and 3a).

Probability maps for each landmark—First, the CNN is used to generate a texture feature vector for each of the landmarks across every patch in the brain (Figure 3b). We then apply the trained classifiers to the feature vectors and generate a separate three-dimensional map for each landmark. These maps report the probability that a given landmark is present at each voxel in the map based only on texture rather than location. The maps for three of the 51 landmarks are illustrated in Figure 3c, where the value of each voxel lies between 0 and 1.

Alignment of reference atlas—We first align the geometrical center of a bounding box for the brainstem of the reference atlas with that of the new brain. This provides a reasonable initial offset for subsequent texture-based alignments. We then simultaneously align the reference atlas to the probability maps for all of the landmarks in the new brain by means of a global affine alignment (Equation 3) (Figure 3d). This transform includes magnification, translation, rotation and shear of the reference atlas; shear corrects for a non-vertical cutting angle. The global alignment is expected to result in a good overlap between the landmarks in the reference atlas and that in the new brain under the constraint that the relative configuration of the landmarks is fixed (Figure 3d). Anatomical information is imposed since the relative position of all landmarks is stable and constrains the probability maps to the correct landmark as false positive results are ignored (Figure 3d).

We next compute a set of individual rigid transforms (Equation 4) that capture the independent variation of each landmark in the new brain (Figure 3e,f). The final fit of each landmark may be verified, and corrected, by human intervention. Figure 3g,h shows the final fit of the reference atlas to the new brain, superimposed on the Nissl stained sections.

The global alignment was formulated to maximize the spatial correlation between the reference atlas and the texture scores for all landmarks at coinciding voxels (Equation 3), while the local alignment maximizes the correlation between the reference atlas and the texture scores for each landmark (Equation 4). To make the local alignment of individual landmarks more robust, the region surrounding the structure was considered in addition to the structure itself. Further, the covariances in centroid positionthat are stored in the reference atlas place landmark-specific constraints on deviations from the nominal position along each axis.

Accuracy and confidence of the alignment of a new brain to the atlas—
Accurate quantification of the position of a landmark is critical for comparisons across brains. We evaluated the automatic alignment of new brains relative to the reference atlas in four ways. The two-dimensional delineations are reconstructed in three dimensions. First, the accuracy of the annotation on the initially annotated brains was assessed by measuring the overlap between the boundaries drawn by the experts and those assigned by our procedure. A simple metric is the fractional overlap, given by the Jaccard index (Equation 6), of the three-dimensional landmarks in new brains with those in the aligned atlas. As an average over 153 landmarks, we achieved a median Jaccard index of 0.61 after the individual

alignments (Equation 4) compared to an index of 0.45 after just the global alignment (Equation 3) (Supplementary Figure 11a).

The second assessment made use of labeling specific landmarks by cell-type specific expression of fluorescent protein. Given the prominence of motor nuclei in the brainstem and the general tight clustering of somata within motor nuclei, we made use of transgenic mice (3 animals) that expressed tdTomato fluorescent protein (FP) that was driven by the promotor for choline acetyltransferase (ChAT) ('Raw' in Figure 4a,b). Motor nuclei that expressed tdTomato FP were manually delineated in images of individual sections using our annotation tool (Supplemental Figure 8).

The reference atlas was formed from thionin rather than Neurotrace blue labeled brains. Are textured derived from these two stains equivalent? Images based on Neurotrace blue staining can be mapped onto those from thionin staining through matching of intensities (Supplementary Figure 13). This permits the thionin-derived classifiers to be used for detection of landmarks with Neurotrace blue images. Yet greater detection accuracy for landmarks in the Neurotrace-stained brains was achieved by fitting classifiers directly to the texture visualized by Neurotrace labeling. This process uses our annotation tool (Supplemental Figure 8) to fit the reference atlas derived from thionin training brains (Figure 2a–d) to a Neurotrace stained brain. The resulting annotations on the Neurotrace images were used to train a new set of classifiers optimized to Neurotrace textures. Note that this procedure to extend the reference atlas is fast, as one does not manually annotate from scratch, and can be used to accommodate any Nissl-like stained brains.

We compared the ChAT delineation with the aligned reference atlas structures, in terms of centroid error and volume overlap. As an average over 15 motor nuclei, we achieved a median Jaccard index of 0.60 after the individual structure alignment (Supplementary Figure 14a). The error in centroid location is typically about 50 μm, which is a small fraction of the size of a motor nucleus. The difference was systematically larger for the case of the X-th motor nucleus. Interestingly, this difference was traced to a bias in the original annotations that excluded neurons at the rostral pole of the tenth motor nucleus (10N) ('Processed' in Figure 4b); this can be used to refine the reference atlas.

For the third assessment, human verification, we asked two experts to review the automatically generated boundaries in nine new, unannotated brains and manually corrected erroneous boundaries Like the local alignment, the experts were only allowed to translate or rotate a given landmark in three-dimensions. We found that in all cases these operations were sufficient to transform unacceptable annotations into reasonable ones. An average of five corrections, out of 51 landmarks, was made on each of nine brains for a 10% false positive rate (Supplementary Figure 12). Of note, less than ten minutes was required for a human to correct the annotations for an entire brain using our annotation tool (Supplemental Figure 8). This is approximately 200-times less than the 30 hours for the initial annotation.

The fourth assessment quantified the confidence of the calculated alignment between the centroids of landmarks in the reference atlas and a new brain. Our procedure was based on the amplitude and width of estimated maxima the global (Equation 3) and local (Equation 4)

alignment objective functions. We quantify the significance of the fit in terms of a z-score, which relates the maximum of these functions relative to their mean in units of standard deviation. For the global alignment of nine new brains, we achieve a median z-score of 2.2 across all landmarks for adjustments in a neighborhood of 50 μm in radius. For local alignments, 90 % of 612 alignments achieved a z-score higher than 1.0, with a median z-score of 1.5 (Supplementary Figure 15a,b). In a companion measurement, the width of the peak of the alignment function was characterized by the Hessian matrix of the z-scores computed at the peak of the distribution. Within the coordinate frame for each landmark, this leads to lower and upper bounds of 66 μm and 193 μm until the z-scores drop to zero, i.e., chance (Supplementary Figure 15c–f).

Update atlas and compute variability in alignment across all brains—We now turn to the variability of the position of landmarks across brains. This measure will consist of the natural biological variability as well as any residual variability from errors in our annotation and our automated procedures. Thus the variability serves as an upper bound on biological variability as well as on our ability to gauge significance in the overlap of labels across brains.

We updated the centroids of the reference atlas with each new brain. We quantified the variability with respect to the updated centroid of each landmark across twelve new brains. This provides a measure of the deviation of every landmark from the sample means (Figure 4d). We observe that some landmark structures are non-isotropic in their variability. For example, the variability of spinal trigeminal nucleus caudalis (Sp5C) is predominantly along the medial-lateral axis, while that of the substantia nigra reticulata (SNr) is primarily along the dorsal-ventral axis (Figure 4d).

As a population over all landmarks and all three axes, the sample-averaged root-mean-square standard deviation is 160 ±40 μm (Figure 4c, Supplemental Figure 16a). This is greater than the typical error in estimating centroids, which is based on comparing the aligned reference atlas against the ground truth, i.e., motonuclei deduced from ChAT (Figure 4a–c) and annotations by experts on the basis of thionin cytoarchitecture (Figure 4c). This suggests that the sample-averaged standard deviation is dominated by biological variability. Of particular note, there was no systematic increase in variability along the rostral-caudal axis (Figure 4e), as might occur from poor brain-to-brain fixation. Other axes showed a similar lack of systematic behavior (Supplemental Figure 16b–d).

Deformation fields

The alignments between the landmarks in the reference atlas and those in a new brain are interpolated to generate a global deformation field (Equation 5). This yields a set of deformation vectors for every location in the tissue sections of the new brain that maps to a location in the atlas (Supplemental Figure 18). This is used to map markers located between landmarks and provides the means to compare the locations of markers, e.g., labeled cells and their projections, across different brains.

Alignment of neuronal projections

As a first example of the utility of automated alignment, we identify the three-dimensional spatial distribution of orofacial premotor neurons labeled with a retrograde viral tracer. Pseudorabies virus that expresses green FP was injected into the masseter muscle, which is responsible for jaw closure. The animal was sacrificed and perfused 86-hours after the injection; at this time all pre-motor neurons and some pre^2-motor are expected to be labeled [23]. We observe extensive labeling of presynaptic populations throughout the brainstem and hypothalamus (Figure 5ai–aiii), yet labeling of trigeminal motor (5N) neurons only on the ipsilateral side, as expected (Figure 5aiv). Known premotor populations were labeled in diverse primary sensory nuclei, e.g., mesencephalic and spinal trigeminal nuclei, the nucleus of the solitary tract, the medial vestibular nucleus, the parvocellular, intermediate, gigantocellular, lateral paragigantocellular regions of the reticular formation, the pontine nucleus and the superior colliculus. These observations replicate known connectivity [24–28]. Yet they further provide the first threedimensional map of trigeminal premotor locations. Additional label in presumed pre^2-motor structures include the central amygdala, the zona incerta, the hypothalamus, and the periaqueductal grey.

As a second example, we assessed the utility of our texture-based alignment for concatenating labeled neurons across multiple brains onto the same coordinate system. We injected retrograde tracers from motorneurons into either the jaw region of the trigeminal motor nucleus (5N) or the he intrinsic vibrissa protractor muscle region of the facial motor nucleus (7N) in separate animals. Specifically, EnvA-pseudotyped G-coat-protein deleted rabies [29] that coded green FP was injected into the respective motonucleus of transgenic mice that expressed the TVA receptor on motoneurons [30]. The brains were processed and countered stained with Neurotrace blue. Two-channel fluorescent detection was used, with blue light for landmark detection and alignment to the reference atlas with our texturebased classifiers, and green light for detecting the viral label. The sagittal three-dimensional projection illustrates the dispersion and heterogeneity of these populations (Figure 5c.ii); red points are premotor neurons of 5N and green points are premotor neurons of the facial motor nucleus (7N). A close-up of the data reveals a subset of two populations with highly overlapped density in the intermediate reticular formation (IRt) and additional overlap in the parvocellular reticular formation (PCRt).(Figure 5c.iii insert). The accurate alignment of fluorescent tracing data illustrates the power of texture-based classifiers, i.e., approximately 90 μm root-mean-square deviation (Figure 5) compared with an approximately 500 μm overlap (Figure 5c.iii insert). Thus texture-based discrimination provides a measure of confidence in the overlap relative to the brain-tobrain variation in landmark positions.

Discussion

We described a method for aligning brains to an atlas, the central step in mapping, that is based on determining and matching the high-resolution statistics between images (Figure 1). Our central advance is that we use the full spatial resolution of the data set to determine the fine-scale texture of all areas in the brain (Figure 2e). We combine this information with the current anatomical atlas. This allows us to determine an accurate alignment of landmarks in the new brain with those in the atlas (Figure 3). We then use the positions of the landmarks

in the new brain to update the mean and variance of all landmarks in the reference atlas (Figure 4). Our approach enables the robust integration of experimental data from different brains in a standard coordinate system (Figure 5). The integration of non-landmark regions, such as newly discovered populations of functionally labeled neurons, occurs through a process of triangulation. Labeled cell populations are mapped using a calculated deformation field. This links labeled cell position to landmarks in the parent brain which then are collectively aligned in relation to the reference atlas landmarks.

Atlases based on texture versus intensity variations

We automate the detection of brain texture at full resolution in single brains and only then combine results from different brains. This is opposed to the adoption of approaches that average variations in section intensity across brains sections in order to define and align landmarks across different brains. Intensity based atlas building is a necessity for MRI brain atlases as slice images are represented at low resolution in grey-levels [31–35]. Intensity-based low resolution detection methods have also been applied to histological data in part to permit co-registration of intensities of histological brain sections to homologous MR imaged brain slices [36,37]. More recently, intensity based detection schemes have been applied to optical sections [38,39] and discrimination of landmark borders is improved by averaging intensity maps across three-dimensional brain reconstructions [5, 40]. An inherent limitation of intensitybased brain registration pipelines is the requirement for additional routines to connect cellular resolution data to intensity- based voxels, as these exceed the typical size of neurons. Recognition of this issue is evidenced by development of software applications to co-register MRI and histological data at cellular resolution [41,42]. An advantage of texture based registration routines is the compatibility of the aligned landmark positioning with cell-based data sets.

We argue that alignment with images that are smoothed by filtering, or by averaging data from multiple brains, will lead to a loss of information about the boundaries of individual landmarks. To illustrate this point, we show the full resolution Nissl stain and convert it to a smoothed image that blurs the Nissl stained texture to mimic a background intensity image that is not Nissl stained, such as those that feed into the Allen Brain Institute atlas [5]. We focus on the oculomotor (3N) and the hypoglossal (12N) motonuclei (Figure 6a–c). Motonuclei are some of the most discernible landmarks in the brainstem. Yet the boundary for the oculomotor nucleus is more difficult to quantify after smoothing (Figure 6b,d,f) while that for the hypoglossal nucleus is clearly obliterated (Figure 6c,e,g). This demonstrates that smoothing even with texture present sufficiently degrades image so as to make boundary detection of low contrast structures difficult. It reinforces our choice of annotating individual brains and then combining the result for formal statistics, an approach that is a necessity when combining brains with different markers. Lastly, prior methods that rely on clusters of neighboring pixels, so called superpixels, of Nissl cytoarchitecture similarly fail to capture local patterns, including cell shape and arrangement [43].

Another departure from past approaches is that we use multiple expert anatomists to bootstraps the atlas (Figure 2c). Moreover, our approach gains in accuracy from the continued involvement of expert anatomists in a number of ways. First, additional annotation

of new landmarks improves and expands the reference atlas. Second, verification of the alignment of individual landmarks improves the accuracy of the centroids and the accuracy of the variation in that centroid (Figure 4c,d). The system maintains the location of each landmark in each brain, the expansion and shear parameters in the global transformation and relative translation in the local transformation. We use these parameters to update the mean and variance of the centroids of each landmark.(Figure 4c,e). With a sufficient number of new annotations (Figures 1 and 3), the shape of each landmark could be updated as well. Lastly, the incorporation of labels to specific markers, e.g, proteins or message RNA, of cell phenotype will increase the accuracy of the cytoarchitecturally based positioning of selected landmarks (Figure 5b), as well as annotates the cellular composition of those landmarks.

Our system does not require perfect data. Although our data underwent good quality control, there remains considerable variability between different images and different parts of an image in terms of brightness, stain quality and focus quality. We train the texture classifiers using such data, which makes the detection robust to normal variations in image conditions. Thus the alignment proceeds well despite the use of classifiers for some landmarks that may be suboptimal, with many false positives (Figure 3c). The confident detection of the characteristic textures of many structures allows specimen-specific deviations from the current reference atlas to be discovered and contributes to an accurate estimate of the variability for each landmark. Simply, the synergy between the anatomical information of landmark location and textural information present in each landmark is a key strength of the active atlas.

Automatic registration of problematic landmarks can fail with our system. One such circumstance is when a landmark is relocated to a nearby region with similar texture. For example, the right occular motor nucleus (3N) may be incorrectly aligned to the left occular motor nucleus, which is immediately adjacent to it. Registration can fail when a structure is incompletely represented in the images. This could occur for structures that are represented in very few sections, such as the abducens motor nucleus (6N), which is as little approximately 50 μm in extent compared with the 20 μm section thickness. Lastly, registration can fail when the textures are diffuse. For example, trigeminal subregions Sp5I and Sp5O are difficult to locate because the boundaries between the subregions are not clearly defined. This inherent difficulty is reflected in their relatively low classification accuracy compared to other structures. In practice, registration of individual landmarks was rare and is corrected by human verification (Figures 1 and 3).

Amortization of labor

The system we describe is effective in amortizing the time spent by experimentalists. Creation of the initial reference atlas of the landmarks involves a heavy investment of time that makes use of multiple expert neuroanatomists and benefits from a diversity annotations and annotators. The payoff from this investment is that the time spent for verification of the position of landmarks in subsequent brains is relatively modest.

There are three contributions to the amortization of labor. First, the alignment of new brains with the atlas is automatic except for a verification step (Table 2). Second, verification involves moving three-dimensional landmarks through the reconstructed volume of all serial

sections of a new brain. Lastly, the verification steps may be accomplished by less experienced anatomists that those needed for the initial annotation.

In the current version of our system, verification by an astute user takes $\simeq$ 5 minutes across the 51 landmarks in our current brainstem atlas. Typically five of the landmarks required a correction, which takes $\simeq$ 1 minute per landmark, or about 10 minutes total after all verification steps. This time must be compared to $\simeq$ the 10 minutes per landmark, in the initial annotations or nearly ten hours per initial brain. While human verification of the automatic alignment is currently the rate-limiting step, the net throughput is now 60-times greater compared to the initial annotation of a brain.

Special challenges of the brainstem

The brainstem contains twelve discrete, well delineated cranial nerve nuclei that serve as part of our set of landmarks. However, unlike forebrain areas with their laminar structure, there is no apparent long-range order to the organization of neurons in the brainstem. Of particular note, the reticular formations are the site of premotor and pre^{2}motor connections that transform sensory input and descending corticobulbar signals into motor actions and behaviors. However, such reticular areas have few clear cytological boundaries that relate function to anatomical structure. Our automated procedure provides a means to localize labeled cells and projections based on their triangulation to landmarks that respects the underlying variability from brain-to-brain (Figure 5).

Extensions

Our method is applicable to the entire vertebrate brain and to the spinal cord, where the issue of illdefined boundaries is especially acute in the cord. More generally, alignment based on texture can provide the underlying computational engine for mature annotation systems and data bases [3,5,11,31,38,40,44–50]. Refinements to particular steps in the method are readily implemented, such as the use of diffeomorphic metric mapping to prevent tears in the deformation field for large deformations [51]. A second extension is to move cytological imaging beyond the necessity for cryostat sections. The challenge is to achieve Nissl-labeling in bulk tissue; nuclear stains such as DAPI and labels such as NeuN fail to report texture [52]. In principle, Nissl-labeling of the whole brain may be achieved by infusing a fluorescent small molecule that stains Nissl bodies, such as methylene blue or cresyl violet, or by constructing a transgenic mouse that mimics this pattern of staining, such as by fluorescently labeled ribosomes. Natural fluorescence, presumably from molecules in the respiratory chain, appears to be be too low in resolution for texturediscrimination [53], although new label-free methods show promise [54]. The brains may then be optically sectioned through a depth of hundreds of micrometers and then resurfaced by mechanical [53] or optical [55] removal of tissue. While improvements in tissue preparation, staining, and microscopy will always improve the practice of mapping, the current work provides a demonstrated means for automated, high resolution alignment.

Methods

Subjects and sample preparation

The dataset for building the atlas consists of 12 brains of postnatal day 56 (P56) male C57BL/6J mice in which all sections were stained with thionin (Supplemental Table 1). We used an additional eight brains of P56 male C57BL/6J mice (JAX no. 000664), three solely for alternate sections of thionin and Neurotrace blue staining, two for injection of the 152 Bartha strain of pseudorabies at a liter of 1×10^9 particles/mL with Neurotrace blue staining, and three for additional tests. Lastly, we used five brains of ChAT-cre mice (JAX no. 006410), two crossed with the FLEX-TVA mice (JAX no. 024708) with injections of EnvA-pseudotyped glycoprotein-deleted rabies-eGFP at a titer of at a liter of 3×10^7 particles/mL (Salk Institute for Biological Studies Virus Core) and three crossed with the Ai14 reporter (JAX no. 007914). All procedures were approved by Institutional Animal Care and Use Committees at the University of California at San Diego and at Cold Spring Harbor Laboratories.

Each brain was cryosectioned in the sagittal plane and mounted using an improved tape-transfer system [21] to yield a set of high-quality 20 μm-thick sections. The sections were stained, cover slipped, and imaged by either a Hamamatsu NanoZoomer at 0.46 μm/pixel resolution and a digitation depth of 8 bits of a Zeiss AxioScan Z.1 at 0.35 μm/pixel resolution and a digitation depth of 16 bits. For animals injected with pseudorabies only, the expression of green FP was enchanced by labeling with anti-GFP (Novos Biologicals NB600–303) visualized with an Alexa-594 labeled secondary. To reduce memory usage for the current analysis, we used only the portion of the images that contain the brainstem, i.e.,~270 sections cropped to ~20,000 by 15,000 pixels. Lastly, since the thionin stain is largely monochrome, we converted these images to grayscale for subsequent processing.

Alignment of images of the serial sectioned brains

Sections acquired with the tape-transfer system have minimal large-scale distortion. To align all sections, we first downsampled the images by a factor of 32; pixel size = 16 μm. We aligned the sections by computing 2-dimensional rigid transforms between every pair of adjacent sections using *Elastix* [56] with the mutual information as the optimization criteria [57]. The correlation is computed using the grayscale downsampled image for thionin sections and using the Neurotrace blue channel for Neurotrace blue images. We then composed these transforms to align each section to the largest section in the brain. To assess the alignment of sections, we inspected virtual coronal slices of the volume reconstructions (Figure 2b). The good quality is demonstrated by the continuity of fine-scale structures such as the hippocampus. As rigid transforms were sufficient to align the sections well, we did not find the need for using more flexible deformable transforms. Finally, we used the transform matrices derived from downsampled images to compute transform matrices that correspond to the full resolution images and brought the raw images into alignment.

Human annotation

Annotation of apparent structural boundaries was performed by two neuroanatomists on the full resolution sagittal images using an in-house program (Supplementary Figure 8). Manual

boundaries were represented by closed polygons and their vertices were recorded. We manually annotated every sections. On average it took an annotator one minute to draw one boundary and 60 hours to annotate a full brain with the 51 selected structures.

Bootstrapping the reference atlas

We converted each set of annotated images of brain sections to a set of threedimensional binary maps that provide the locations of different pre-averaged landmarks, i.e., the landmark from an individual unannotated brain. The voxel size of the map is 10 μm. First, the manual boundaries within individual sections for each structure were spaced in parallel planes according to the section spacing of 20 μm and interpolated to 10 μm resolution to achieve isotropic voxels. Next, to distinguish voxels that are inside versus outside of the preaveraged landmark, a binary map was formed by filling the voxels in the landmark with a value of one and setting the value of all other voxels to zero. These maps were used to compute the nominal position and the nominal shape, as an average over each set of annotations of a given landmark per hemisphere.

Estimating the center-of-mass of landmarks—First the brains were co-aligned under the same coordinate space. The brain with the largest volume was selected as the target and the other brains were aligned to it. Alignment of two brains began with aligning the mid-sagittal planes, which were estimated by fitting to midway points of the centroids of paired structures. Under this constraint, we found an affine transform that maximizes the correlation between the two sets of binary maps; see Global alignment of a new brain with the reference atlas below. Once all brains were aligned, we computed the mean and covariance matrix of the coordinates of the center-of-mass over all annotated brains; three in the present case. The mean was used as the nominal position of the landmark and the covariance matrix was used to regularize its alignment, as described in the section on Landmark-specific alignment below.

Estimating nominal shapes—To estimate the nominal shape of a landmark, we aligned all instances of the preaveraged landmarks from the individual annotated brains by maximizing the overlap of the pre-averaged landmarks using rigid transforms. A probabilistic average shape was then created by counting the percentage of pre-averaged landmarks that contain each voxel (Supplementary Figure 10c–f). Intuitively, the reference atlas is defined by situating the centroid of each shape at its corresponding nominal position.

Training texture classifiers

Patches of grayscale, fullresolution images serve as inputs to the classifiers. We found that a size around 100 μm, or 224 by 224 pixels, shows both local brain organization and detailed cell shape. Larger patches are also effective (Supplementary Figure 17) but may fail to capture small structures. Patches are collected based on a moving window with a pitch of 32 μm that yielded roughly 40,000 patches per section. Training patches for a certain structure are collected from all sections on which this structure was annotated. A patch is labeled positive if at least three of the four corners are located inside a boundary of this structure (Figure 2e). Similarly, a negative patch must have three corners in the bordering zone of a boundary. The use of negative patches that lie in the boundary region, rather than anywhere

in the image, improves the fine-scale localization of landmarks without impairing the large-scale fit of the reference atlas to a new brain.

We used the Inception-bn CNN [22] (implemented by *MXNet* [58]) to encode the patches. This CNN had previously been trained on a subset of ImageNet, a dataset of 21,000 natural scene images in 1000 categories, and achieved state-of-the-art classification performance. We modified the network to accommodate single-channel input, and used the 1024-dimensional vector that feeds into the last fully-connected layer as features of the patches.

The texture feature vectors were used to train binary logistic regression classifiers (Equation 1), which were implemented by Python *scikit-learn*. Logistic regression assumes a linear prediction model and finds a weight vector that maximizes the likelihood of the input data. Suppose for a given structure, n training patches are used. We denote the feature vector of the i-th patch by x_i and its label by y_i (Figure 2e). The L2-penalized logistic regression minimizes:

$$\sum_{i=1}^{n} \log\left(1 + e^{-\left(y_i \mathbf{x}_i \cdot \mathbf{w} - \theta\right)}\right) + \alpha \|\mathbf{w}\|^2 . \quad (2)$$

The optimal weight vector w and offset θ define the classifier for this landmark.

Automated landmark detection for new unannotated brains

Given a new brain, we applied the full set of classifiers to a moving window on every section. Suppose the feature vector of a patch is x and the weight vector of a particular classifier is w, then the predicted probability is $y = \sigma(\mathbf{x} \cdot \mathbf{w} - \theta)$, where $\sigma(z) = (1 + e^{-z})^{-1}$. For each classifier, the predicted probabilities for all windows on all sections formed a sparse three-dimensional probability map. This was then resampled using cubic interpolation and discretized to create a dense map with voxel size of 16 μm on edge (Figure 3c). The resolution of these volumes was low so that they can be simultaneously loaded into the computer memory as required by the global alignment algorithm.

Global alignment of a new brain with the reference atlas—Alignment of the new brain occurs by correlating the three-dimensional texture scores across all landmarks with the landmarks in the reference atlas (Figure 3d). Specifically, we computed a three-dimensional affine transform that maximizes the total correlation between all pairs of texture probability maps over the entire domain. The affine transform can be represented jointly by a matrix $\mathbf{A} \in \mathbb{R}^{3 \times 3}$ and a shift vector $b \in \mathbb{R}^3$. The transform maps a coordinate x in the reference atlas to another coordinate $\mathbf{Ax} + \mathbf{b}$ in the input brain.

Denote by Φ the set of all landmarks. For a particular landmark r, denote the probability map of the input brain by S_r and that of the atlas by Q_r. Ω_r is a subdomain of the reference atlas that contains the landmark r, as well as the surrounding area. Global alignment was formulated as maximizing the sum:

$$F^g(\mathbf{A}, \mathbf{b}) = \sum_{r \in \Phi} \sum_{\mathbf{x} \in \Omega_r} S_r(\mathbf{Ax} + \mathbf{b})\big(Q_r(\mathbf{x}) - 0.5\big) \quad (3)$$

The optimal $\mathbf{A}$ and b are found by stochastic gradient ascent. At each iteration the Jacobian is computed based on ten thousand randomly sampled voxels from each structure. The adaptive gradient algorithm Adagrad was employed to automatically control the learning rate. Convergence was usually achieved in less than one hundred iterations.

Landmark-specific alignment—After the global affine transform adjusted the pose of the new brain to be roughly the same as that of the reference atlas, we estimated the deviations of different landmarks from their nominal positions. In this case we compute a rigid transform separately for each landmark. The three-dimensional rigid transform for structure r is denoted by $G(\mathbf{x}; \omega_r, \mathbf{u}_r) = \mathbf{R}(\omega_r)\mathbf{x} + \mathbf{u}_r$, where $\mathbf{u}_r \in \mathbb{R}^3$ is the shift vector and $\mathbf{R}(\omega_r) \in \mathbb{R}^{3 \times 3}$ is a rotation matrix parametrized by the Euler vector $\omega_r \in \mathbb{R}^3$.

For a given structure r, the objective function F_r^l only involves the probability map corresponding to this particular landmark and only concerns the subdomain around it. A regularization term is added to penalize large deviations; this term is based on the position covariance matrix $\mathbf{C}_r$ stored with the reference atlas, so that deviations in different directions are penalized differently. We maximize:

$$F_r^l(\mathbf{u}_r, \omega_r) = \sum_{\mathbf{x} \in \Omega_r} S_r\big(G(\mathbf{x}; \omega_r, \mathbf{u}_r)\big)\big(Q_r(\mathbf{x}) - 0.5\big) - \beta\mathbf{u}_r^T\mathbf{C}_r^{-1}\mathbf{u}_r \quad (4)$$

where β is the regularization weight; $\beta = 0.01$ in our experiments. Optimization used gradient ascent on the logarithmic mapping of Lie group SO(3). Convergence was usually achieved in 30 iterations.

Deformation field

In order to transform the positions of molecular markers between the landmarks, we interpolated the local transforms using the centroids of the landmarks as control points. This yielded a deformation field that was defined for every point in the reference atlas (Supplemental Figure 18), within and outside all landmarks. For location $\mathbf{x}$, the deformation vector is expressed as:

$$D(\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \sum_r b\big(\|\mathbf{x} - \mathbf{c}_r\|\big) G(\mathbf{x}; \omega_r, \mathbf{u}_r) \quad (5)$$

where $\mathbf{c}_r$ is the centroid of landmark r after alignment, $Z(\mathbf{x}) = \sum_r b\big(\|\mathbf{x} - \mathbf{c}_r\|\big)$, and b is a radial basis function that computes the influence of a control point based on distance. We used $b(d) = 1/d^2$.

Evaluation of alignment accuracy for brains with ground truth evaluation

After computing the global transform and the landmark-specific transform for each landmark, we warped each probability map of the reference model to fit the input brain using the composition of both transforms. The warped atlas maps can be sliced at the position of particular sections and thresholded to generate structure boundaries on the original images (Figure 3).

In manually annotated brains, the landmark structures derived from automatic alignment were compared to manual annotations, using the isosurface for a probability of $p = 0.5$. For each pair of boundaries for a given structure on the same image, we computed the centroid-to-centroid distance in three-dimensions and the Jaccard index between the associated three-dimensions binary masks. The Jaccard index, ranging between 0 and 1, measures the overlap of two binary masks A and B, and is defined by:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}. \quad (6)$$

Evaluating alignment confidence for brains lacking ground truth evaluation

In addition to accuracy, we evaluated the confidence of each alignment. Specifically, we quantified the height and width of the found objective function maximum (Equation 3 and 4).

Peak height—The value of the maximum was normalized by the mean and standard deviation of the values in a neighborhood around the maximum, similar to the computation of a z-score. The neighborhood includes translations of ± 50 μm in three directions and rotations of ± 15 degrees around three axes.

Peak width—We computed the Hessian matrix of the objective function at the maximum with respect to translations in three directions. Based on the eigenvalues and eigenvectors of the Hessian, we derived the most certain and the least certain translation directions which were not necessarily paraxial. In addition, we computed for each of the directions a "margin", defined as the amount of deviation from the maximum along the given direction that the $z$score drops to one.

Normalization of fluorescent images

In our dataset, the brightfield thionin-stained sections are imaged at 8-bit and the fluorescent Neurotrace blue-stained sections are 16-bit. While thionin staining is fairly uniform, the fluorescence intensity for Neurotrace staining has sufficient variability between different sections, or different parts of the same section, to confound texture-based learning. We mitigated this issue with an adaptive procedure that uses a moving window to high pass filter as well as normalize the data. We first choose moderately-sized windows that are evenly spaced across image. For each window we compute a linear correction factor to make the pixel values have zero mean and unit standard deviation. Correction factors across adjacent sections are interpolated to generate a smooth correction surface of factors for each pixel. In

detail, 2 mm by 2 mm windows are taken across an image with 1.2 mm even spacing. For each window the mean μ and standard deviation σ of the pixel values are computed. Bilinear interpolation of each correction factor of all window centers gives the correction factors for every pixel. The new intensity value of a pixel x is $v'(x) = -\mu(x)v(x) + 1/\sigma(x)$. This normalization step eliminates the variability in fluorescent intensity that is irrelevant to texture. It is crucial for the successful learning of texture classifiers and the accurate detection on new brain section images.

Training a separate set of classifiers for Neurotrace blue images

We utilized the reference atlas to reduce annotation time. Using the in-house program (Supplementary Figure 8), two neuroanatomists manually shifted and rotated the probabilistic landmark structures defined in the reference atlas to best fit the images. The probability level at which to extract the isosurface was hand-picked for each structure. Once the annotations in the form of twodimensional structure boundaries are obtained, the same procedure for training thionin detectors was used to train this new set of classifiers specific to Neurotrace images. Intensity normalized Neurotrace blue channels are used for training and testing.

Data availability

All raw data is publicly available. It may be downloaded, with a listing of files found in Supplemental Table 1 and at https://github.com/ActiveBrainAtlas/MouseBrainAtlas/blob/master/doc/Brain_stack_directories.md, through the Amazon Web Service Storage S3 at the bucket named mousebrainatlas-rawdata.

Code availability

All analysis was done following the algorithms detailed in this Methods section. The code is written in Python and is available at https://github.com/ActiveBrainAtlas/MouseBrainAtlas through the GNU General Public License (GPL). Organization of the code is in a ReadMe.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

[1]. Roland PE & Zilles K Brain atlases: A new research tool. Trends in Neurosciences 17, 458–467 (1994). [PubMed: 7531886]

[2]. Jones EG, Stone JM & Karten HJ Highresolution digital brain atlases: A hubble telescope for the brain. Annals of the New York Academy of Sciences 1225 S1, E147E159 (2011). [PubMed: 21599693]

[3]. MacKenzie-Graham A et al. A multimodal, multidimensional atlas of the c57bl/ 6j mouse brain. Journal of Anatomy 204, 93102 (2004).

[4]. Majka P & Wojcik DK Possuma framework for three-dimensional reconstruction of brain images from serial sections. Neuroinformatics 14, 265278 (2016).

[5]. Kuan L et al. Neuroinformatics of the allen mouse brain connectivity atlas. Methods 73, 4–17 (2015). [PubMed: 25536338]

[6]. Pauli WM, Nil AN & Tyszka JM A highresolution probabilistic in vivo atlas of human subcortical brain nuclei. Scientific Data 5, 180063 EP (2018). [PubMed: 29664465]

[7]. Toga AW et al. Postmortem cryosectioning as an anatomic reference for human brain mapping. Computerized Medical Imaging and Graphics 21 (1997).

[8]. Swanson LW & Bota M Foundational model of structural connectivity in the nervous system with a schema for wiring diagrams, connectome, and basic plan architecture. Proceedings of the National Academy of Sciences USA 107, 20610–20617 (2010).

[9]. Jones EG Viewpoint: The core and matrix of thalamic organization. Neuroscience 85, 331–345 (1998). [PubMed: 9622234]

[10]. Braitenberg V On the Texture of Brains, An introduction to Neuroanatomy for the Cybernetically Minded (Springer Verlag, 1977).

[11]. Gong H et al. High-throughput dual-colour precision imaging for brain-wide connectome with cytoarchitectonic landmarks at the cellular level. Nature Communication 7, 12142 (2016).

[12]. Economo MN et al. A platform for brain-wide imaging and reconstruction of individual neurons. eLIFE 5, e10566 (2016). [PubMed: 26796534]

[13]. Richardson DS & Lichtman JW Clarifying tissue clearing. Cell 162, 246257 (2015).

[14]. Wilt BA et al. Advances in light microscopy for neuroscience. Annual Review of Neurosciences 32, 435–506 (2009).

[15]. Gray PA Transcriptional factors define the neuroanatomical organization of the medullary reticular formation. Frontiers in Neuroanatomy 7, 1–21 (2013). [PubMed: 23423949]

[16]. McElvain LE et al. Circuits in the rodent brainstem that control whisking in concert with other orofacial motor actions. Neuroscience 368, 152–170 (2018). [PubMed: 28843993]

[17]. Chiang A-S et al. Three-dimensional reconstruction of brain-wide wiring networks in drosophila at single-cell resolution. Current Biology 21, 1–11 (2011). [PubMed: 21129968]

[18]. Peng H et al. Brainaligner: 3d registration atlases of drosophila brains. Nature Methods 8, 493–498 (2011). [PubMed: 21532582]

[19]. Ronneberger O et al. Vibe-z: A framework for 3d virtual colocalization analysis in zebrafish larval brains. Nature Methods 9, 735–742 (2012). [PubMed: 22706672]

[20]. Randlett O et al. Whole-brain activity mapping onto a zebrafish brain atlas. Nature methods 12, 1039–1046 (2015). [PubMed: 26778924]

[21]. Pinskiy V et al. High-throughput method of whole-brain sectioning, using the tape-transfer technique. PLoS ONE 10 (2015).

[22]. Ioffe S & Szegedy C Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015).

[23]. Fay RA & Norgren R Identification of rat brainstem multisynaptic connections to the oral motor nuclei using pseudorabies virus. i. masticatory muscle motor systems. Brain Research Brain Research Reviews 25, 255–275 (1997). [PubMed: 9495558]

[24]. Yasui Y et al. Non-dopaminergic neurons in the substantia nigra project to the reticular formation around the trigeminal motor nucleus in the rat. Brain Research 585, 361–366 (1992). [PubMed: 1380874]

[25]. Li Y, Takada M, Kaneko T & Mizuno N Premotor neurons for trigeminal motor nucleus neurons innervating the jaw?closing and jaw?opening muscles: Differential differential in the lower brainstem of the rat. Journal of Comparative Neurology 365, 563–579 (1995).

[26]. Mizuno N et al. A light and electron microscopic study of premotor neurons for the trigeminal motor nucleus. Journal of Comparative Neurology 215, 290–298 (1983). [PubMed: 6304157]

[27]. Travers JB & Norgen R Afferent projections to the oral motor nuclei in the rat. Journal of Comparative Neurology 220, 280–298 (1983). [PubMed: 6315785]

[28]. Stanek E, Rodriguez E, Zhao S, Han BX & Wang F Supratrigeminal bilaterally projecting neurons maintain basal tone and enable bilateral phasic activation of jaw-closing muscles. Journal of Neuroscience 36, 7663–7675 (2016). [PubMed: 27445144]

[29]. Wickersham IR, Finke S, Conzelmann K-K & Callaway EM Retrograde neuronal tracing with a deletion-mutant rabies virus. Nature Methods 4, 47–49 (2007). [PubMed: 17179932]

[30]. Takatoh J et al. New modules are added to vibrissal premotor circuitry with the emergence of exploratory whisking. Neuron 77, 346–360 (2013). [PubMed: 23352170]

[31]. Johnson GA et al. Waxholm space: An image-based reference for coordinating mouse brain research. Neuroimage 53, 365–372 (2010). [PubMed: 20600960]

[32]. Roland PE et al. Human brain atlas: For high-resolution functional and anatomical mapping. Human Brain Mapping 1, 173184 (1994).

[33]. Pollack JD, Wu D-Y & Satterlee JS Molecular neuroanatomy: a generation of progress. Trends in Neurosciences 37, 106–123 (2014). [PubMed: 24388609]

[34]. Gonzlez-Vill S et al. A review on brain structures segmentation in magnetic resonance imaging. Artificial Inteligence in Medicine 73, 45–69 (2016).

[35]. Papp EA, Leergaard TB, Calabrese E & Johnson GA Waxholm space atlas of the sprague dawley rat brain. Neuroimage 97, 374–386 (2014). [PubMed: 24726336]

[36]. MacKenzie-Graham A et al. The informatics of a c57bl/6j mouse brain atlas. Neuroinformatics 1, 397–410 (2003). [PubMed: 15043223]

[37]. Yushkevich PA et al. Using mri to build a 3d reference atlas of the mouse brain from histology images. In Proceedings of the International Society of Magnetic Resonance in Medicine, vol. 13 (2005).

[38]. Oh SW et al. A mesoscale connectome of the mouse brain. Nature 508, 201–214 (2014).

[39]. Renier N et al. Mapping of brain activity by automated volume analysis of immediate early genes. Cell 165, 1789–1802 (2016). [PubMed: 27238021]

[40]. Feng D et al. Exploration and visualization of connectivity in the adult mouse brain. Methods 73, 9097 (2015).

[41]. Lau C et al. Exploration and visualization of gene expression with neuroanatomy in the adult mouse brain. BMC Bioinformatics 9, 153 (2008). [PubMed: 18366675]

[42]. Dempsey B et al. Mapping and analysis of the connectome of sympathetic premotor neurons in the rostralventrolateral medulla of the rat using a volumetric brain atlas. Frontiers of Neural Circuits 11, 9 (2017).

[43]. Senyukova OV, Lukin AS & Vetro DP Automated atlas-based segmentation of nissl-stained mouse brain slices. Programming and Computer Software 37, 245–251 (2011).

[44]. Amunts K & Zilles K Architectonic mapping of the human brain beyond brodmann. Neuron 88, 1086–1107 (2015). [PubMed: 26687219]

[45]. Frth D et al. An interactive framework for whole-brain maps at cellular resolution. Nature Neuroscience 21, 139149 (2018).

[46]. Bakker R, Tiesinga P & Ktter R The scalable brain atlas: Instant web-based access to public. Neuroinformatics 13, 353366 (2013).

[47]. Zingg B et al. Neural networks of the mouse neocortex. Cell 156, 10961111 (2014).

[48]. Ng L et al. An anatomic gene expression atlas of the adult mouse brain. Nature Neuroscience 12, 356–362 (2009). [PubMed: 19219037]

[49]. Mazziotta J et al. A probabilistic atlas and reference system for the human brain: International consortium for brain mapping (icbm). Philosophical Transactions of the Royal Society of London B: Biological Sciences 356, 1293–1322 (2001). [PubMed: 11545704]

[50]. Parekh R & Ascoli GA Neuronal morphology goes digital: A research hub for cellular and system neuroscience. Neuron 77, 1017–1038 (2013). [PubMed: 23522039]

[51]. Miller MI, Beg MF, Ceritoglu C & Stark C Increasing the power of functional maps of the medial temporal lobe by using large deformation diffeomorphic metric mapping. Proceedings of the National Academy of Sciences USA 102, 9685–9690 (2005).

[52]. Tsai PS et al. All-optical histology using ultrashort laser pulses. Neuron 39, 27–41 (2003). [PubMed: 12848930]

[53]. Ragan T et al. Serial two-photon tomography for automated ex vivo mouse brain imaging. Nature Methods 9, 255–258 (2012). [PubMed: 22245809]

[54]. Ren J, Choi H, Chung K & Bouma BE Label-free volumetric optical imaging of intact murine brains. Scientific Reports 7, 46306 (2017). [PubMed: 28401897]

[55]. Tsai PS et al. All-optical histology using ultrashort laser pulses. Neuron 39, 27–41 (2003). [PubMed: 12848930]

[56]. Klein S, Staring M, Murphy K, Viergever MA & Pluim JP Elastix: a toolbox for intensity-based medical image registration. Medical Imaging, IEEE Transactions on 29, 196–205 (2010).

[57]. Maes F, Collignon A, Vandermeulen D, Marchal G & Suetens P Multimodality image registration by maximization of mutual information. Medical Imaging, IEEE Transactions on 16, 187198 (1997).

[58]. Chen T et al. Mxnet: A flexible and efficient machine learning library for heterogeneous distributed systems. arXiv preprint arXiv:1512.01274 (2015).
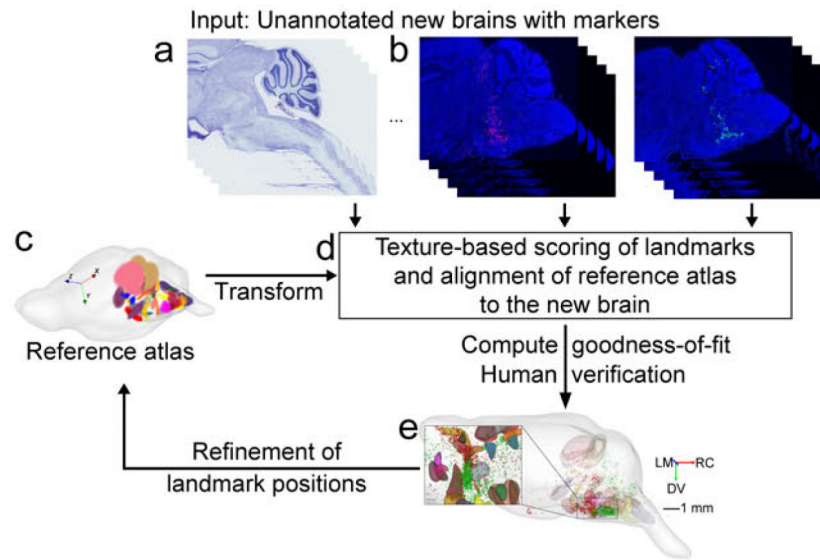
Figure 1. Overarching structure of an automated atlas.
(**a,b**) Inputs to the system are histological sections from a new brain, in these examples horizontal sections of a mouse brain with Nissl staining, exemplified by thionin-stained cells for brightfield data (panel a) and Neurotrace blue stained cells for fluorescent data (panel b); the brains in panel b are colabeled with red and green tracers, respectively. (**c**) The reference atlas, in this case with only brainstem landmarks. (**d**) Computational steps involve the scoring of features, texture in our case, for the alignment of the new brain with the reference atlas. Human experts may then review the alignment and make corrections if necessary to the position of specific landmarks. (**e**) Alignment of two brains to the reference atlas to illustrate the power of the automated atlas. One brain contained G-rabies injected into the jaw region of the trigeminal motor nucleus, while the other contained G rabies injected into the vibrissa region of the facial motor nucleus; in both cases motor and premotor neurons are labeled by the expression of green fluorescent protein. The combined data set shows an overlap of premotor neurons, red points for jaw and green points for vibrissa, in the parvocellular region of the reticular formation. The aligned new brains are further used to refine the landmark positions of the reference atlas.
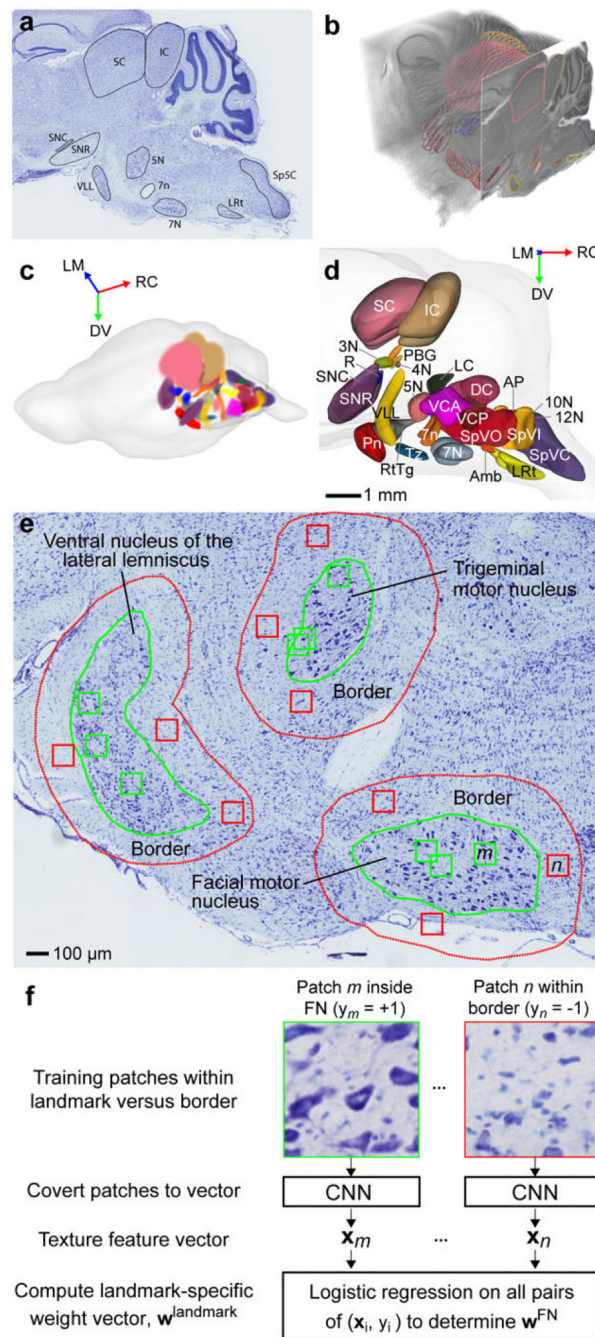
Figure 2. Workflow for training the atlas, which consists of annotating brain sections followed by computation.

The input for training was a set of sagittally cut sections of the entire mouse brain, at a thickness of 20 μm, that were stained with thionin and imaged in bright-field a 0.5 μm resolution. (**a**) Expert annotation of landmarks and their boundaries in one section. (**b**) Resliced, three-dimensional view of a stack of successive sections with annotated boundaries. (**c**) The initial reference atlas that bootstrapped from the expert annotation. Fuzzy boundaries highlight the probabilistic nature of the shapes of the landmarks as an

average across annotations and annotators. The directions are dorsal-ventral (D-V), rostral-caudal (R-C), and lateral-medial (L-M). (**d**) The 28 structures in one hemisphere in the current reference atlas (Table 1). Surfaces correspond to $p^{\text{landmark}} = 0.5$. (**e**) Three representative image patches in an annotated section that are used to train the texture-based binary classifiers (Eq. 1). Patches inside the landmark are extracted from the interior of boundaries (green boxes) and tagged as positive, i.e., $y_m = +1$, while patches in a 200 μm wide moat that surrounds the landmark (red boxes) are tagged as negative, i.e., $y_m = -1$). (**f**) Training of the classifier for the example of the facial motor nucleus (FN). Each training patch is converted to a feature vector, e.g., x_m, using a convolutional neural network (CNN) with fixed weights. The classifier, f^{FN}, is a function of the weight vector, w^{FN}.
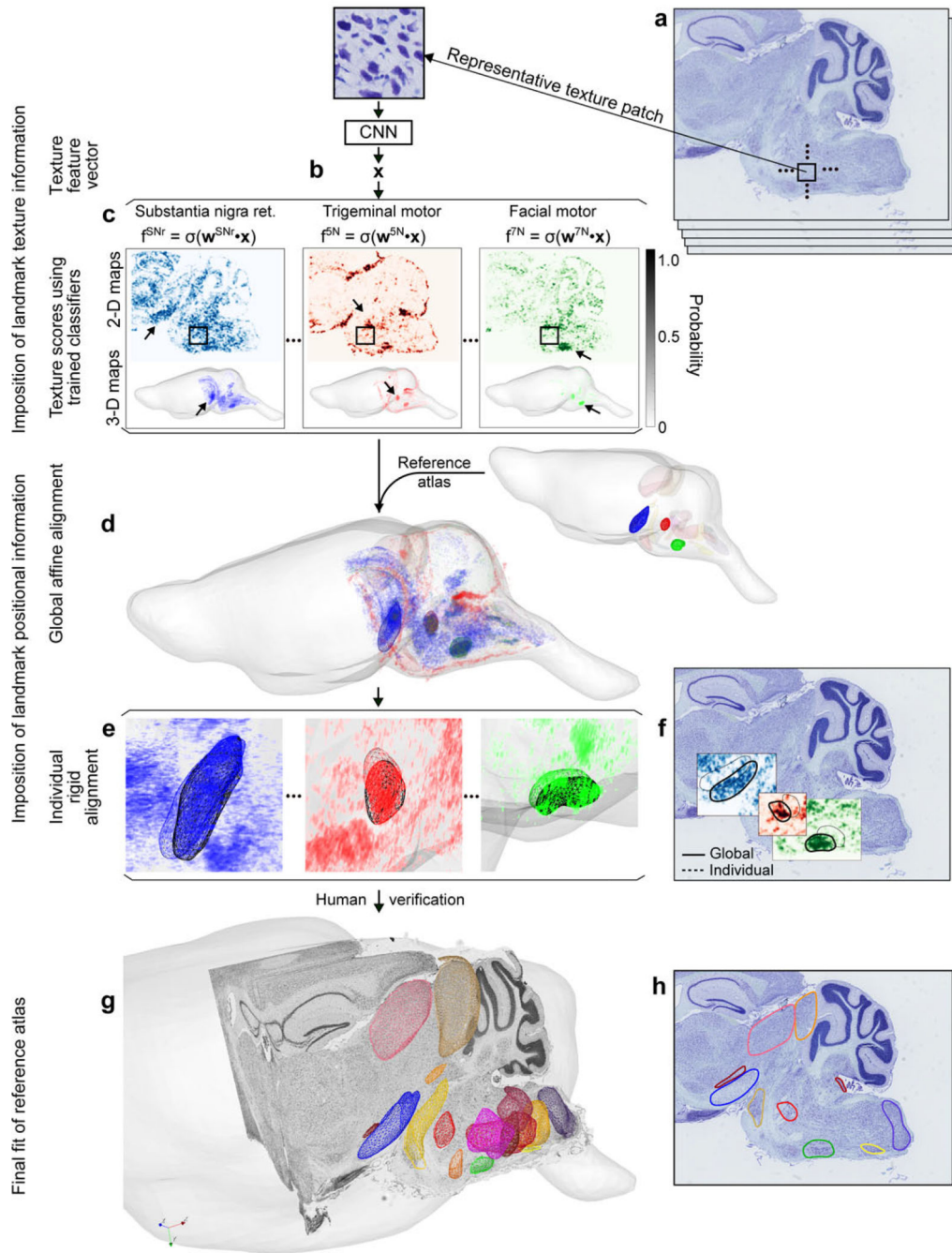
Figure 3. Workflow to align a new brain with the current reference atlas.
The input has the same Nissl-stain as the training set. (**a**) An unannotated set of Nissl-stained sections from a new brain. (**b**) One example patch that is passed through a CNN to be converted into a texture feature vector **X**. (**c**) Example of three of the 51 texture-based classifiers for each landmark that are applied to all texture feature vectors in the brain patches across the entire brain. This results in a probability map for each landmark, illustrated here for one section and throughout the brain. Note that the raw data is downsample along the R-C and D-V directions directions by a factor of 32 to achieve

isotropic pixels of 20 μ on edge. (**d**) The atlas after global affine alignment to the probability maps for all landmarks. (**e, f**) Local alignment between individual landmarks in the new brain with those in the atlas in three-dimensions (panel e) and for one section superimposed on the classifier scores (panel f). The thin mesh is the initial position and thick mesh is the final position. Contours are cross-sections of the transformed nominal shapes at $p^{\text{landmark}} =$ 0.5. (**g**) Illustration of the final aligned result. Greyscale image volume is the reconstruction of the Nissl sections. Colored structures are the transformed reference atlas. (**g**) Contour lines from the aligned reference atlas overlaid on the section in panel a.
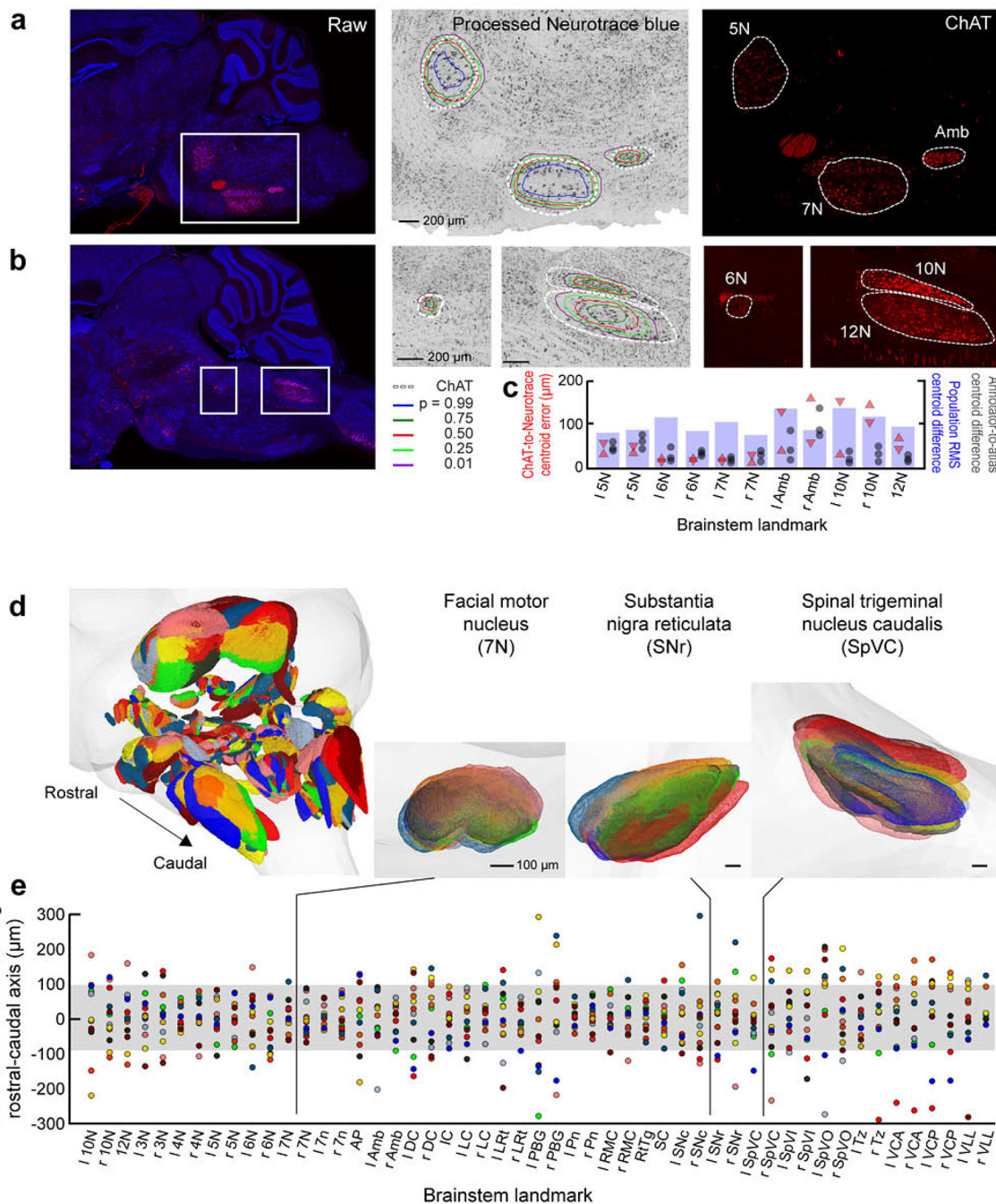
Figure 4. Reliability and variability in estimates of landmark position for new brains.
(**a,b**) An assessment based on comparison of landmark positions found with our texture-based classifiers, using a brain in which Nissl bodies were labeled with Neurotrace Blue, to that found with the centroids for ChAT-tdTomato FP labeled brains. Eleven motor nuclei were compared. The plots show the raw. two channel data followed by close-up views of selected motor nuclei in the Neurotrace blue and tdTomato FP channels. (**c**) Compendium of the difference, in three dimensions, between centroids found from the Neurotrace channel, which reports Nissl bodies, compared to that from the ChAT channel for two brains (red

triangles). Also shown is the variation of the positions noted by each of three human annotations from the mean position in the reference atlas (gray circles). Lastly, we plot the root-mean-square (RMS) variations between centroids across 12 brains (blue bars). (**d**) Shift in position of the aligned landmarks from nine new brains from the centroids of landmarks in the initial reference atlas. Different new brains are represented by different colors. We show both the full brainstem and for three example landmarks. (**e**) A compendium of the shift in rostral-caudal position of the centroids for for all landmark across all new brains.
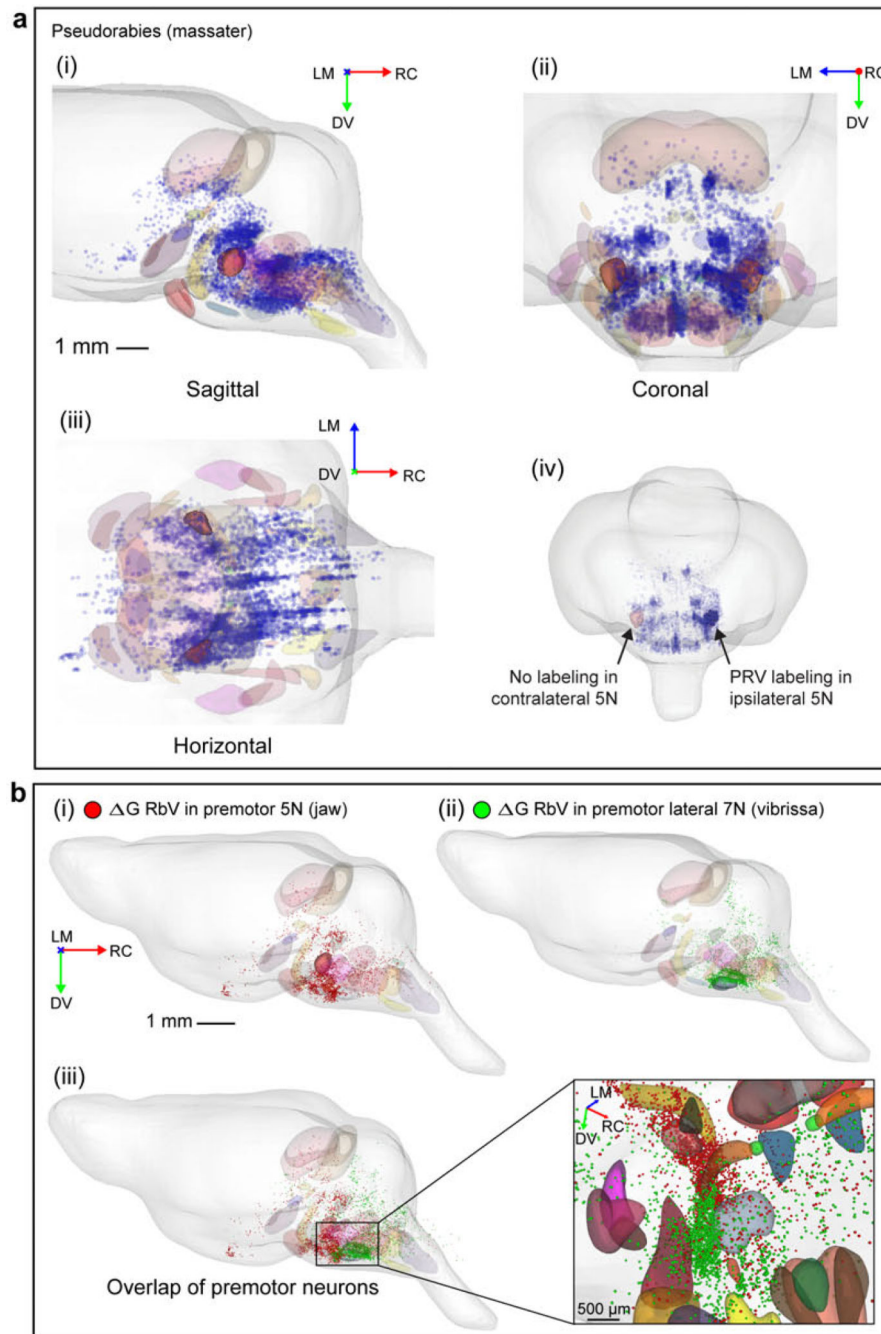
Figure 5. Application of the texture-based alignment to fluorescent imaging within and across brains.

(**a**) Visualization of the labeling motor and premotor inputs to the jaw muscle across all three planes (subpanels i through iii). The masseter muscle was injected with pseudorabies (PRV) that expressed green FP and visualized with a Neurotrace background stain. The PRV labeled cells were manually annotated and aligned with the reference atlas. Note the widespread, bihemispheric inputs and, critically, the absence of labeling from the contralateral motor nucleus (subpanel iv). (**b**) Visualization of the labeling of different

populations of premotor neurons in separate brains with overlapped density in the parvocellular reticular formation (PCRt). We labeled the premotor neurons of the jaw region of the trigeminal motor nucleus (5N) using G-protein deleted rabies (G-RV) that expressed green FP (red points; subpanel i) and the premotor neurons of the vibrissa region of the facial motor nucleus (7N) using the same construct (green points subpanel ii). Premotor neurons predominantly overlap in a border area of the reticular formations IRt and PCRt (subpanel iii). The insert shows a magnified view of the overlap of the two premotor populations (Figure 1).
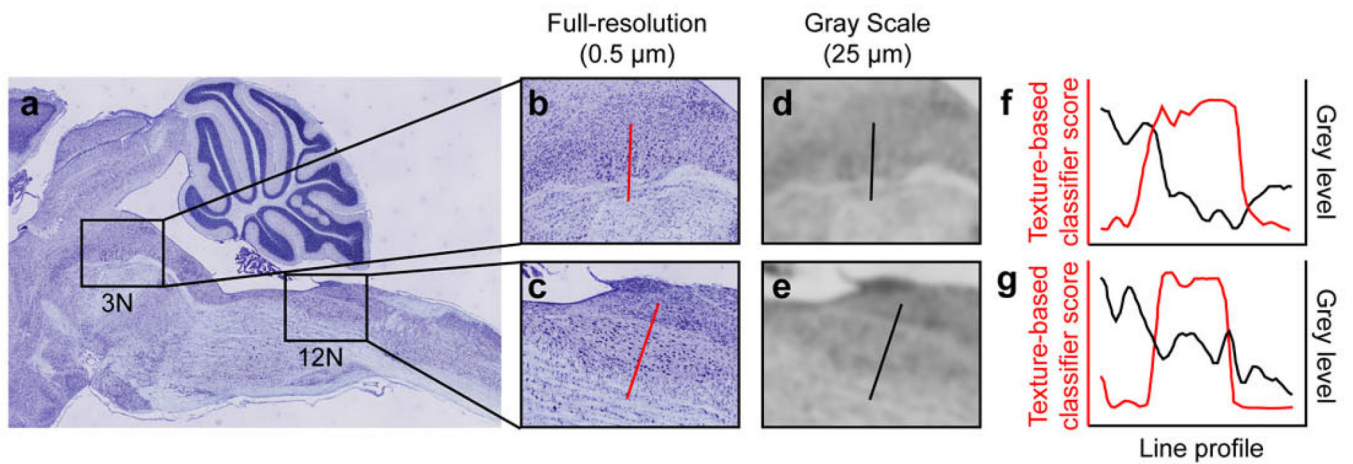
Figure 6. Defining landmark based on texture versus gray scale.
The gray-scale was formed by downsampling from 0.5 to 25 μm resolution. (**a**) Image of the hindbrain and midbrain of a single section stained with thionin. The boxes contain two motor nuclei: the oculomotor nucleus (3N) and the hypoglossal nucleus (12N). (**b,c**) The area around each nucleus at higher magnification and a resolution of 0.5 μm. The nucleus appears distinct from its surrounding area. (**d,e**) The same areas as in panels b and c after downsampling to 25 μm of resolution. The semi-homogeneous intensity within a landmark is required by traditional alignment methods, but precision is sacrificed by the blurred boundaries. (**f,g**) Comparison of the probability of a landmark from the texture-based classifier versus the value of gray levels along the lines across the nuclei in panels b to e. All values are normalized. Texture-detection results in a steeper plateau to yield a precise alignment.

Table 1:

Landmark structures in active atlas

Symbol	Name	R/L fused
3N	Occulomotor nucleus	N
4N	Trochlea nucleus	N
5N	Trigeminal motor nucleus	N
6N	Abducens nucleus	N
7N	Facial motor nucleus	N
7n	Facial nerve	N
10N	Dorsal nucleus of vagus nerve	N
12N	Hypoglossal nucleus	Y
Amb	Nucleus ambiggus	N
AP	Area postrema	Y
DC	Dorsal cochlea nucleus	N
LRt	Lateral reticular nucleus	N
LC	Locus corelus	N
IC	Inferior colliculus	Y
VCA	Ventral cochlea nucleus, anterior	N
VCP	Ventral cochlea nucleus, posterior	N
VLL	Ventral lateral lemniscus	N
PBG	Parabigeminal nucleus	N
Pn	Pontine grey	N
R	Red nucleus	N
RtTg	Reticulotegmental nucleus	Y
SC	Superior colliculus	Y
Sp5C	Spinal-trigeminal nucleus, caudalis	N
Sp5I	Spinal-trigeminal nucleus, interpolaris	N
Sp5O	Spinal-trigeminal nucleus, oralis	N
SNR	Substantia niagra, reticular	N
SNC	Substantia niagra, compact	N
Tz	Nucleus of trapezoidal body	N

Table 2:

Typical computation times for a new brain

Step	Time
Intra-stack registration using downsampled images	0.7 hours
Transform and crop raw images	7.5 hours
Compute features	1.5 hours (8 GPUs) 5.0 hours (1 GPU)
Generate probability maps	0.7 hours
Registration	0.5 hours