nature methods

ARTICLES

https://doi.org/10.1038/s41592-019-0328-8

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*] and David Kleinfeld[2,5*]

Brain atlases enable the mapping of labeled cells and projections from different brains onto a standard coordinate system. We address two issues in the construction and use of atlases. First, expert neuroanatomists ascertain the fine-scale pattern of brain tissue, the 'texture' formed by cellular organization, to define cytoarchitectural borders. We automate the processes of localizing landmark structures and alignment of brains to a reference atlas using machine learning and training data derived from expert annotations. Second, we construct an atlas that is active; that is, augmented 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.

Brain atlases provide a depository for the ever-expanding studies of neuron wiring and function[1,2]. The navigability of any atlas depends on demarcation of regional boundaries, or landmarks. The standard for brain atlas construction uses sets of landmarks that are shared across brains to define a reference atlas[3–5] and 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 resonance imaging (MRI) reference brain atlases is based on low resolution images where boundaries are determined from large shifts in gray levels; recent brain atlases have adopted these standards to retain a compatible modality with a three-dimensional reference space dictated by MRI of a representative brain[5]. This approach limits the types of landmark that can be used for navigation, as regions with subtle boundaries are not recognized; for these reasons, numerous small brainstem structures in mice have not been absorbed into a standardized reference atlas. Additionally, fixed reference atlases do not incorporate the expected variance in brain regions of subject mice brains, even though brains of inbred mice can differ in the structural characteristics of neurons within a common region[11]. 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. 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 higher-resolution spatial information acquired by 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 machine-generated detection of multiple landmarks in the new brain using the texture classifiers (Fig. 1a–d). Final verification is performed by a human. Thus the software system aligns 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 atlas also uses landmarks to triangulate regions with subtle ill-defined borders and then coaligns such regions across separate brains with both high precision and known uncertainty (Fig. 1e).

We apply our approach to the murine brainstem, that is, the hindbrain and midbrain, across a cohort of mice. The brainstem is a challenging region to map. 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 neuronal populations in subregions with, at best, subtle borders[15]. These premotor regions, mainly in the extensive reticular formation of the brainstem, are important in regulation of

[1]Department of Computer Science and Engineering, University of California, San Diego, CA, USA. [2]Department of Physics, University of California, San Diego, CA, USA. [3]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA. [4]Department of Neurosciences, University of California School of Medicine, San Diego, CA, USA. [5]Section of Neurobiology, University of California, San Diego, CA, USA. *e-mail: yfreund@eng.ucsd.edu; dk@physics.ucsd.edu

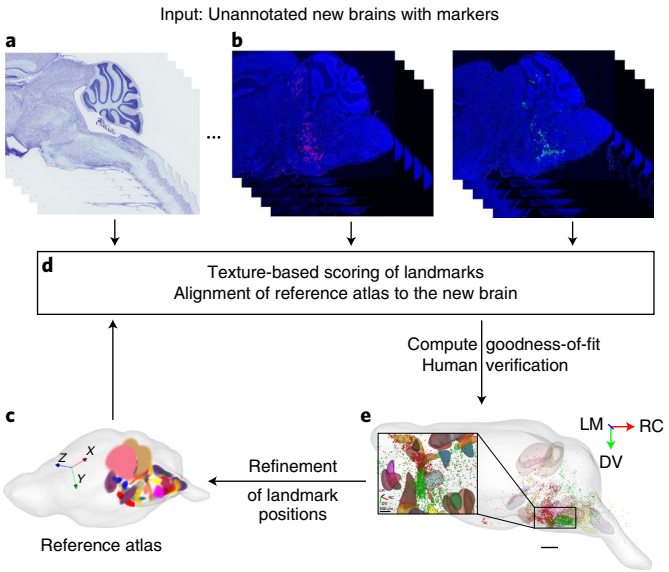

Fig. 1 | Structure of an automated atlas. a,**b**, Inputs to the alignment system are histological sections from a new brain. In these examples sagittal sections of a mouse brain stained for Nissl bodies, either with thionin for brightfield data (**a**; one of 12 brains) or Neurotrace blue for fluorescent data (**b**; one of two brains). **c**, The reference atlas, in this case with only brainstem landmarks. **d**, Computational steps involve the scoring of texture 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 ΔG-rabies-green fluorescent protein (GFP) injected into the jaw region of the trigeminal motor nucleus, while the other contained ΔG-rabies-GFP injected into the vibrissa region of the facial motor nucleus; in both cases motor and premotor neurons are labeled by the expression of GFP. The aligned brains show an overlap of premotor neurons, red points for jaw and green points for vibrissa, in the parvocellular region of the reticular formation and are further used to refine the landmark positions of the reference atlas. The directions are dorsal–ventral (DV), rostral–caudal (RC), and lateral–medial (LM). Scale bar, 1 mm.

brainstem output functions that range from breathing to orofacial sensorimotor behaviors[16]. The failure to form a reliable atlas of this region has stifled comparisons of studies across brains and from different laboratories. Our approach is general and 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 to identify landmarks for the alignment of brains. We sectioned brains from P56, male C57BL/6 mice in a sagittal plane on a cryostat and maximized the quality of our histological sections through the use of an improved tape-transfer method[21]. This procedure minimizes 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, that is, ribosomal and messenger 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 (Fig. 2a,b and Supplementary Fig. 1). The process is aided by a display of the annotation that numerically constructs views of the data in the two alternate planes in real-time (Supplementary Fig. 2). The annotated data forms a training set for our texture-based classifiers and captures the location and approximate shape of each of the landmarks to bootstrap the reference atlas. In practice, our experts contoured around each of 51 landmarks across three brains (Fig. 2c,d); these correspond to 28 different structures. Note that the right and left sides of five structures that border the midline, the hypoglossal nucleus, inferior and superior colliculus, area postrema and reticulotegmental nucleus, were fused into a single landmark.

Training structure-specific texture classifiers. We divide the image of each brain section into overlapping square patches that are large enough to contain many cells but small enough so that each landmark is tiled by many patches. The training patches are 100 µm on the edges with a pitch of 30 µm. Patches within an annotated landmark are labeled positive, that is, $y_m = +1$ for the mth patch, while patches in a boundary region that surrounds the landmark are labeled negative; that is, $y_m = -1$ (Fig. 2e). The textural information of each image patch is encoded as a set of numbers, called a feature vector and denoted by $\mathbf{x}_m$. To perform encoding, we used a convolutional neural network (CNN) with fixed weights, that is, the blue channel only of the Inception-BN[22], which was trained on natural scenes. The rich internal filters appear to effectively represent histological textures in terms of a 1,024-dimensional vector that defines $\mathbf{x}_m$, so that each patch is represented by the pair $(\mathbf{x}_m, y_m)$.

We used supervised learning to create the texture-based classifiers, one for each landmark and denoted f^{landmark}. Bilateral landmarks share classifiers. 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 (Fig. 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, $\mathbf{w}^{\text{landmark}}$ and an offset θ^{landmark} and 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, that is,

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

and is a number between 0 and 1. The weight vectors and offsets are found by maximizing the likelihood of the training data. No averaging of either images or textures is involved. The complete set of classifiers, parameterized by $\mathbf{w}^{\text{landmark}}$ and θ^{landmark}, enable us to score a new brain for the probability, $f^{\text{landmark}}(\mathbf{x}_n)$, that the nth 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 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 (Supplementary Fig. 3) with a mean 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 volumetric annotations that jointly constitute a labeled volume for each annotated brain. The labeled volumes of all annotated brains were co-aligned and the mean and covariance of the coordinates of the centroid for each landmark were computed (Supplementary Fig. 4a,b) We further derived 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 (Supplementary Fig. 4c–f). We labeled the regions that are included in the annotations of all brains by $p^{\text{landmark}} = 1$, while regions that were incorporated by only

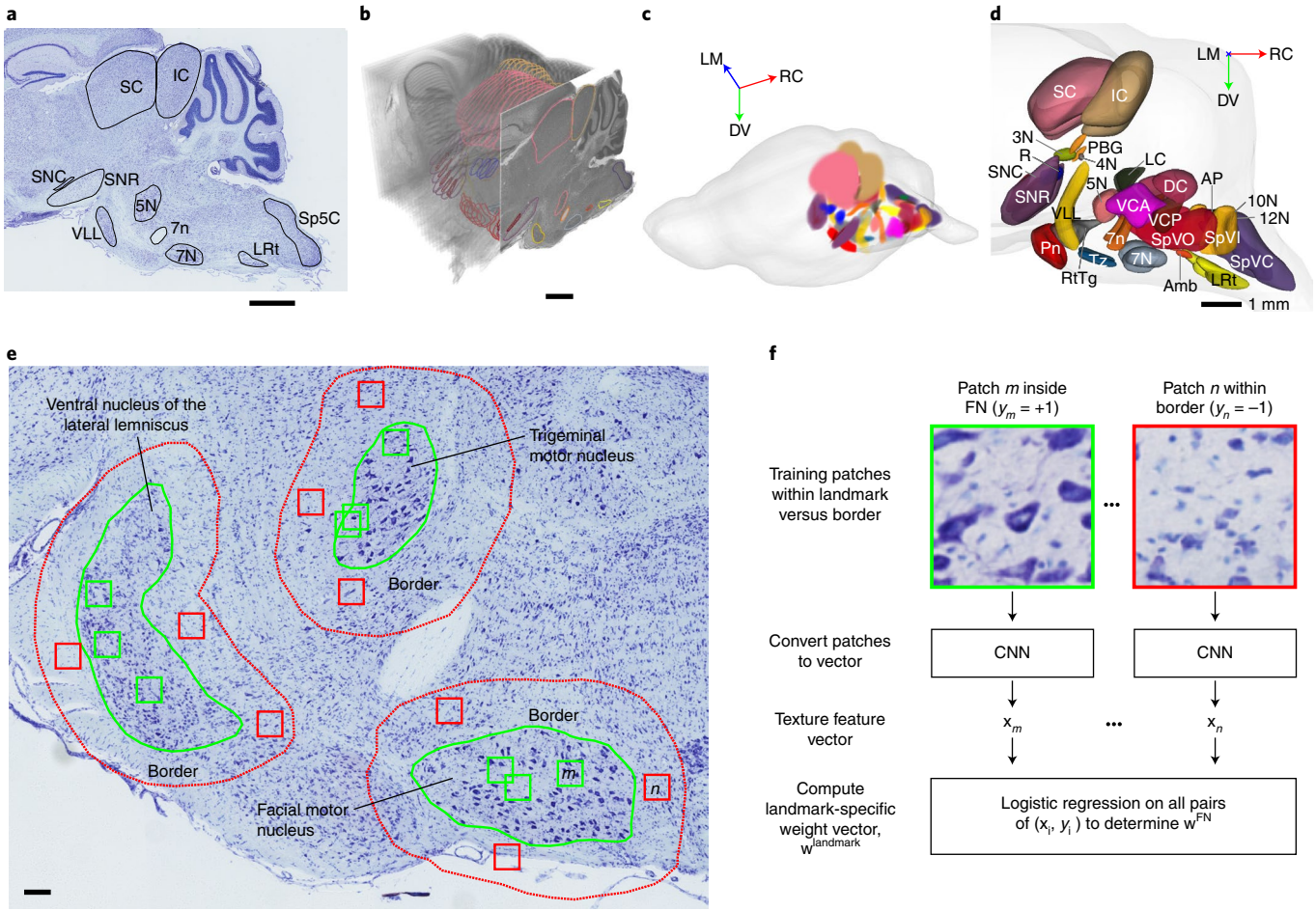

Fig. 2 | Workflow for training the atlas, which consists of annotating brain sections followed by computation. The input for training was a set of sagittally cut sections of the entire mouse brain at a thicknesses of 20 µm that were stained with thionin and imaged in brightfield at a 0.46 µm resolution. **a**, Expert annotation of landmarks and their boundaries in one section. **b**, Three-dimensional view of a stack of successive sections with annotated boundaries. **c**, The initial reference atlas that bootstrapped from the expert annotation (three brains). Fuzzy boundaries highlight the probabilistic nature of the shapes of the landmarks as an average across annotations and annotators. The directions are dorsal–ventral (DV), rostral–caudal (RC) and lateral–medial (LM). **d**, The 28 structures in the current reference atlas. Surfaces correspond to $p^{\text{landmark}} = 0.5$. Abbreviations: 3N, oculomotor nucleus; 4N, trochlear nucleus; 5N, trigeminal motor nucleus; 6N, abducens nucleus; 7N, facial motor nucleus; 7n, facial nerve; 10N, dorsal nucleus of vagus nerve; 12N, hypoglossal nucleus; Amb, nucleus ambiguus; AP, area postrema; DC, dorsal cochlear nucleus; LRt, lateral reticular nucleus; LC, locus coeruleus; IC, inferior colliculus; VCA, ventral cochlear nucleus, anterior; VCP, ventral cochlear nucleus, posterior; VLL, ventral nucleus of the lateral lemniscus; PBG, parabigeminal nucleus; Pn, pontine gray; R, red nucleus; RtTg, reticulotegmental nucleus; SC, superior colliculus; Sp5C, spinal-trigeminal nucleus, caudalis; Sp5I, spinal-trigeminal nucleus, interpolaris; Sp5O, spinal-trigeminal nucleus, oralis; SNR, substantia nigra, reticulata; SNC, substantia nigra, compacta and Tz, nucleus of trapezoidal body. Scale bar, 1 mm in **a**,**b**,**d**. **e**, Representative image patches in an annotated section that are used to train the texture-based binary classifiers (equation 1). Patches inside the landmark are extracted from the interior of boundaries (green boxes) and tagged as positive, that is, $y_m = +1$, while patches in a 200 µm wide moat that surrounds the landmark (red boxes) are tagged as negative; that is, $y_m = -1$). Scale bar, 100 µm. **f**, Training of the classifier for the example of the facial motor nucleus (7N). Each training patch is converted to a feature vector, for example, $\mathbf{x}_m$, using a convolutional neural network (CNN) with fixed weights. The classifier, f^{FN}, is a function of the weight vector, $\mathbf{w}^{\text{FN}}$.

a fraction of the annotations had $p^{\text{landmark}} < 1$. The combination of the average shapes and mean centroids of all the landmarks gave rise to the initial probabilistic reference atlas (Fig. 2c,d and Supplementary Fig. 4d–f).

Automated alignment of a new brain with reference atlas. We used the trained classifiers and the reference atlas (Fig. 2c,d) to align a new serially sectioned brain with the reference atlas. We first used additional thionin counterstained sections to test the accuracy and reproducibility of our approach (Figs. 1a and 3a).

Probability maps for each landmark. First, the CNN was used to generate a texture feature vector for each of the landmarks across every patch in the brain (Fig. 3b). We then applied the trained classifiers to the feature vectors and generated a separate three-dimensional map for each landmark. These maps reported 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 landmarks are illustrated in Fig. 3c, where the value of each voxel lies between 0 and 1.

Alignment of reference atlas. We first aligned the geometrical center of a bounding box for the brainstem of the reference atlas with that of the new brain. This provided an initial offset for subsequent texture-based alignments. We then simultaneously aligned the reference atlas to the probability maps for all of the landmarks in the new brain by a global affine alignment (equation 3) (Fig. 3d). This transform included magnification, translation, rotation and shear of the

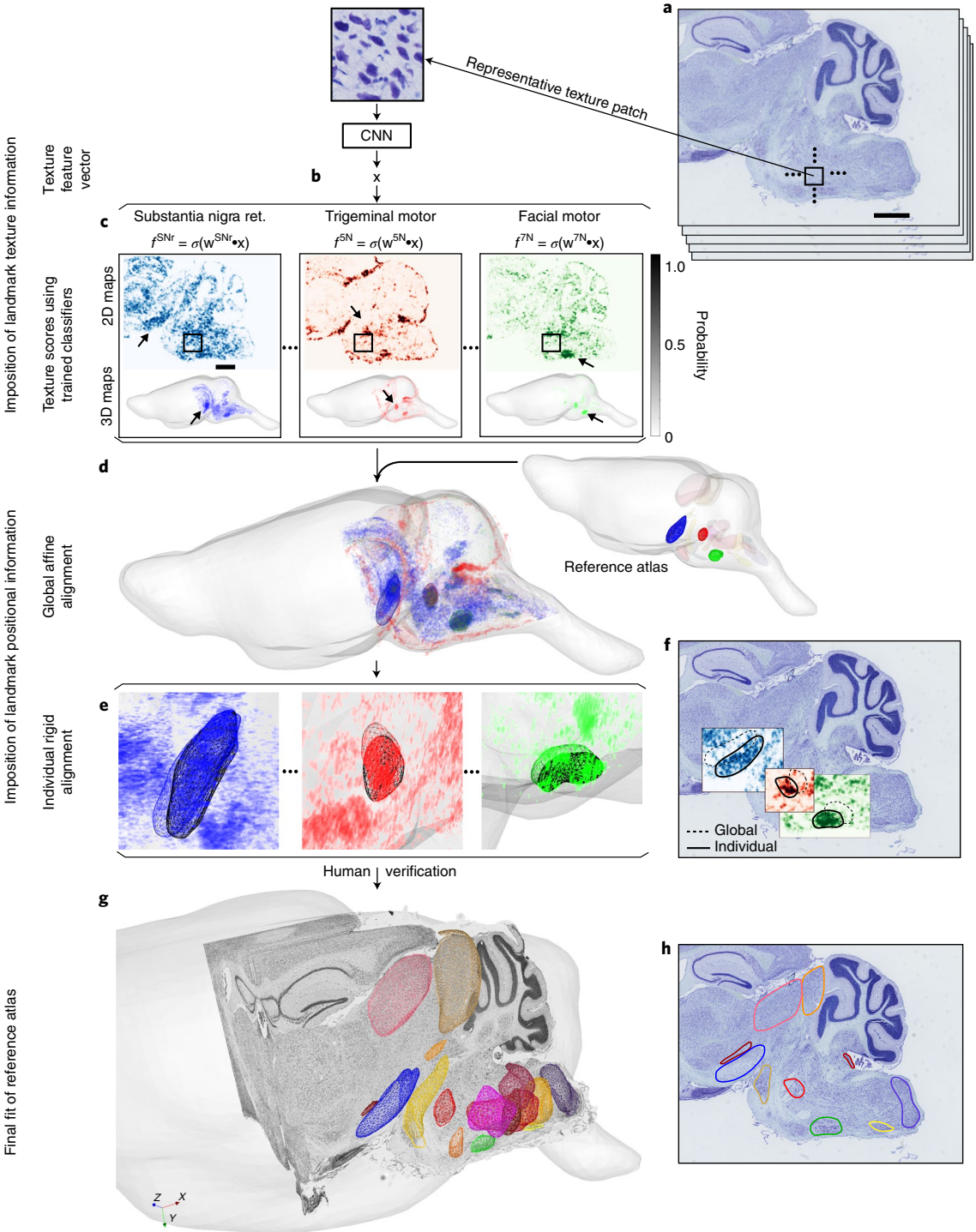

Fig. 3 | Workflow to align a new brain with the current reference atlas. The input has the same Nissl stain as the training set. **a**, An unannotated set of Nissl-stained sections from a new brain. Size of texture box is exaggerated. Scale bar, 1 mm. **b**, One example patch that is passed through a CNN to be converted into a texture feature vector **x**. **c**, Example of three 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. Size of texture box is exaggerated. 2D, two-dimensional; 3D, three-dimensional. Scale bar, 1 mm. **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 (**e**) and for one section superimposed on the classifier scores (**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 $p^{landmark} = 0.5$. **g**, Illustration of the final aligned result. Grayscale 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 **a**.

reference atlas; shear corrected for a non-vertical cutting angle. The global alignment was 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 was fixed (Fig. 3d). Anatomical information was imposed since the relative position of all landmarks was stable and constrained the

probability maps to the correct landmark as false positive results were ignored (Fig. 3d).

We next computed a set of individual rigid transforms (equation 4) that captured the independent variation of each landmark in the new brain (Fig. 3e,f). The final fit of each landmark could 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.

We formulated the global alignment 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, we considered the region surrounding the structure in addition to the structure itself. Further, the covariances in centroid position that were stored in the reference atlas placed landmark-specific constraints on deviations from the nominal position along each axis, i.e., small variances permitted only small deviations by the local alignment.

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 were reconstructed in three dimensions. First, we assessed the accuracy of the annotation on the initially annotated brains 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 5), 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 Fig. 5).

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 used transgenic mice (two animals) that expressed Cre driven by the promoter for choline acetyltransferase (ChAT) that were crossed with a tdTomato fluorescent protein reporter line ('Raw' in Fig. 4a,b). We manually delineated motor nuclei that expressed tdTomato fluorescent protein in images of individual sections using our annotation tool (Supplementary Fig. 2).

The reference atlas was formed from thionin rather than Neurotrace blue labeled brains. Are textures derived from these two stains equivalent? It might not be expected that the textures obtained from Neurotrace blue fluorescent Nissl and thionin brightfield Nissl would be identical, especially as the mechanism of staining by the two dyes may not be equivalent. Images based on Neurotrace blue staining can be mapped onto those from thionin staining through matching of intensities (Supplementary Fig. 6). This permitted the thionin-derived classifiers to be used for detection of landmarks with Neurotrace blue images. Yet, we achieved greater detection accuracy for landmarks in the Neurotrace-stained brains by fitting classifiers directly to the texture visualized by Neurotrace labeling. This process used our annotation tool (Supplementary Fig. 2) to fit the reference atlas derived from thionin training brains (Fig. 2a–d) to a Neurotrace-stained brain. We used the resulting annotations on the Neurotrace images 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 it 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 Fig. 7). The error in centroid location is typically about 50 µm, which is a small fraction of the size of a motor nucleus (Fig. 4c). The difference was systematically larger for the case of the tenth motor nucleus (10N). We traced this difference to a bias in the original annotations that excluded neurons at the rostral pole of the tenth motor nucleus ('processed' in Fig. 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 to manually correct erroneous boundaries. As for 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 Fig. 8). Note that a human required less than 10 min to correct the annotations for an entire brain using our annotation tool (Supplementary Fig. 2). This is approximately 200 times less than the 30 h 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 for 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 Fig. 9a,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 and 193 µm until the z-scores drop to zero; that is, chance (Supplementary Fig. 9c–f).

Update atlas and compute variability in alignment across all brains. The variability of the position of landmarks across brains consists 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 12 new brains. This provides a measure of the deviation of every landmark from the sample means (Fig. 4d). 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 (Fig. 4d).

As a population over all landmarks and all three axes, the sample-averaged r.m.s. standard deviation is 160 ± 40 µm (Fig. 4c and Supplementary Fig. 10a). This is greater than the typical error in estimating centroids, which is based on comparing the aligned reference atlas against the ground truth, that is, motonuclei deduced from expression of ChAT (Fig. 4a–c), and annotations by experts on the basis of thionin cytoarchitecture (Fig. 4c). This suggests that the sample-averaged standard deviation is dominated by biological variability. Of note, there was no systematic increase in variability along the rostral-caudal axis (Fig. 4e), as might occur from poor brain-to-brain fixation. Other axes showed a similar lack of systematic behavior (Supplementary Fig. 10b–d).

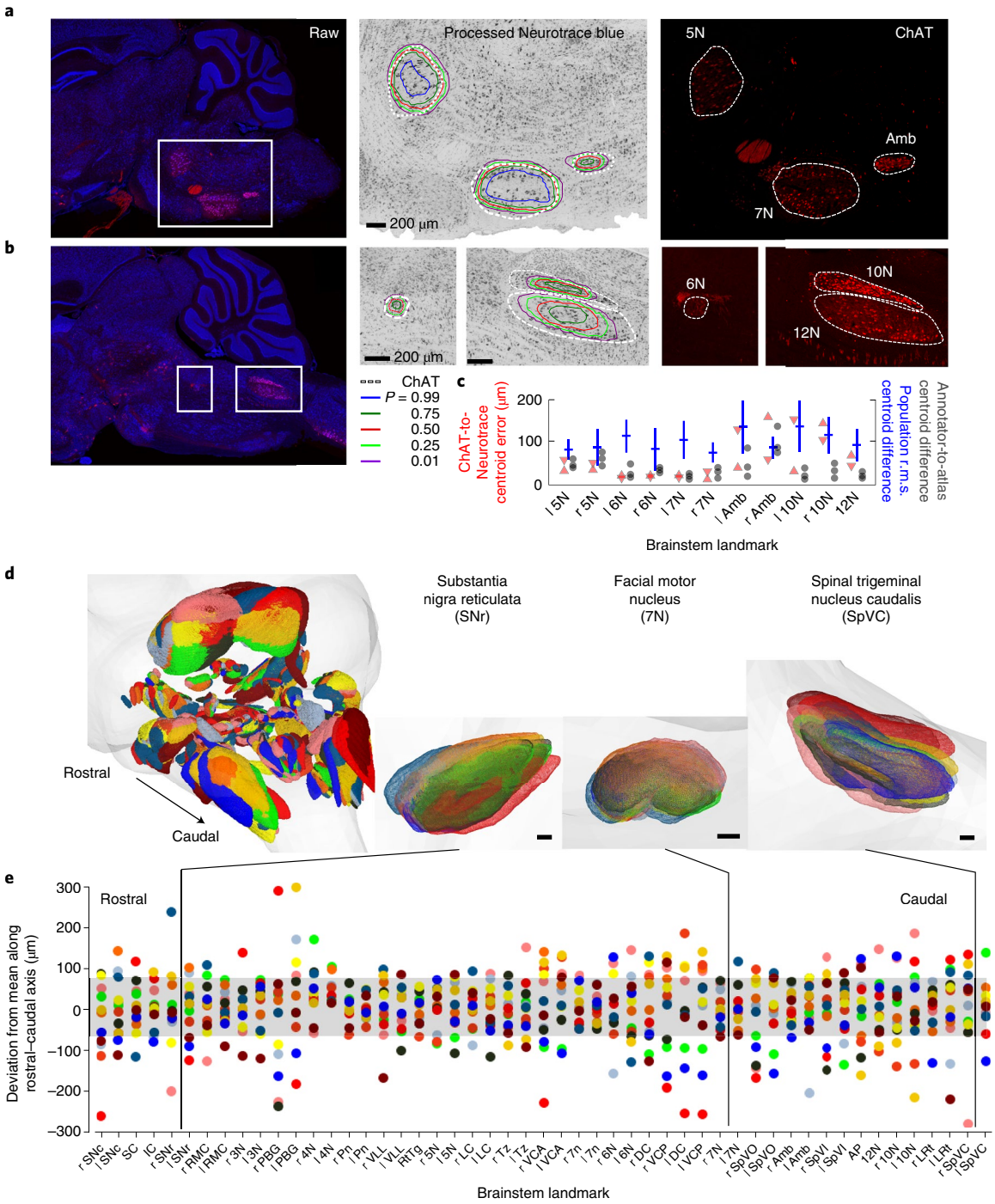

Fig. 4 | Reliability and variability in estimates of landmark position for new brains. a,**b**, Comparison of landmark positions found with our texture-based classifiers, using a brain in which Nissl bodies were labeled with Neurotrace Blue, to that found with the centroids for ChAT-tdTomato fluorescent protein labeled brains. Eleven motor nuclei were compared in two brains. The plots show the raw, two-channel data (left) for one mouse and close-up views of selected motor nuclei in the Neurotrace blue (center) and tdTomato fluorescent protein (right) channels. Scale bars, 200 µm. **c**, The difference, in three dimensions, between centroids found from the Neurotrace channel compared to those from the ChAT channel for two brains (red triangles). Variation of the positions noted by each of three human annotations from the mean position in the reference atlas (gray circles). Root-mean-square (r.m.s.) variations between centroids across 12 brains (blue crosses); the heights represent the standard deviation of the estimated r.m.s. value. **d**, Shift in position of the aligned landmarks from nine new brains from the centroids of landmarks in the initial reference atlas. Different brains are represented by different colors. We show both the full brainstem and three example landmarks. Scale bars, 100 µm. **e**, Shift in rostral–caudal position of the centroids for all landmark across all new brains. A listing of all abbreviations can be found in the legend for Fig. 2.

Deformation fields. We interpolated the alignments between the landmarks in the reference atlas and those in a new brain to generate a global deformation field (equation 4). This yielded a set of deformation vectors for every location in the tissue sections of the new brain that mapped to a location in the atlas (Supplementary Fig. 11). This was used to map markers located between landmarks

and provided the means to compare the locations of markers, for example, labeled cells and their projections, across different brains.

Alignment of neuronal projections. As a first example of the use of automated alignment, we identified the three-dimensional spatial distribution of orofacial premotor neurons labeled with a retrograde viral tracer. We injected pseudorabies virus (PRV) that expressed GFP into the masseter muscle, which is responsible for jaw closure. The animal was euthanized and perfused 86 h after the injection; at this time all pre-motor neurons and some pre^2-motor were expected to be labeled[23]. We observed extensive labeling of presynaptic populations throughout the brainstem and hypothalamus (Fig. 5a–c), yet there was labeling of trigeminal motor (5N) neurons only on the ipsilateral side (Fig. 5d). Pre-motor populations were labeled in diverse primary sensory nuclei, for example, 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, replicating known connectivity[24–28]. This further provides the first three-dimensional map of trigeminal premotor locations. Additional labels in presumed pre^2-motor structures included the central amygdala, the zona incerta, the hypothalamus and the periaqueductal gray.

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 motor neurons into either the jaw region of the trigeminal motor nucleus (5N) or the intrinsic vibrissa protractor muscle region of the facial motor nucleus (7N) in separate animals. Specifically, we injected EnvA-pseudotyped G-coat-protein deleted rabies[29] that coded GFP into the respective motonucleus of transgenic mice that expressed the TVA receptor on motoneurons[30] (Fig. 5e,f). We processed and counterstained the brains with Neurotrace blue. We used two-channel fluorescent detection, with blue light for landmark detection and alignment to the reference atlas with our texture-based classifiers, and green light for detecting the viral label. The sagittal three-dimensional projection illustrates the dispersion and heterogeneity of these populations (Fig. 5g); 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 and additional overlap in the parvocellular reticular formation (inset of Fig. 5g). The accurate alignment of fluorescent tracing data illustrates the power of texture-based classifiers, that is, approximately 70 µm r.m.s.d. (Fig. 5) compared with an approximately 500 µm overlap (Fig. 5g). Thus, texture-based discrimination provides a measure of confidence in the overlap of premotor populations relative to the brain-to-brain variation in landmark positions.

Discussion

We have described an automated method for aligning brains to an atlas, the central step in mapping, which is based on determining and matching the high-resolution statistics between images of brain cytoarchitecture (Fig. 1). We detected brain texture at full resolution in single brains and only then combined results from different brains. This is opposed to the adoption of approaches that average variations in section intensity across brain sections 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 gray 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 MRI 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 intensity-based 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.

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 showed the full-resolution Nissl stain and converted it to a smoothed image that blurred the Nissl-stained texture to mimic a background intensity image that was not Nissl-stained, such as those that feed into the Allen Brain Institute atlas[5]. We focused on the oculomotor (3N) and the hypoglossal (12N) motonuclei (Fig. 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 (Fig. 6b,d,e) while that for the hypoglossal nucleus is clearly obliterated (Fig. 6c,f,g). This demonstrates that smoothing, even with texture present, sufficiently degrades the 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. Last, previous methods that relied on clusters of neighboring pixels, the so-called 'superpixels' of Nissl cytoarchitecture, similarly failed to capture local patterns, including cell shape and arrangement[43].

Another departure from past approaches is that we used multiple expert anatomists to bootstrap the atlas (Fig. 2c). Moreover, our approach gained in accuracy from the continued involvement of expert anatomists. First, additional annotation of new landmarks improved and expanded the reference atlas. Second, verification of the alignment of individual landmarks improved the accuracy of the centroids and the accuracy of the variation in that centroid (Fig. 4c,d). The system maintained 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 used these parameters to update the mean and variance of the centroids of each landmark (Fig. 4c,e). With a sufficient number of new annotations (Figs. 1 and 3), the shape of each landmark could be updated as well. Last, the incorporation of labels to specific markers, for example, proteins or messenger RNA, of cell phenotypes increased the accuracy of the cytoarchitecturally based positioning of selected landmarks (Fig. 5a–c), as well as annotating the cellular composition of those landmarks.

Our system does not require perfect data. Although our data underwent good quality control, there remained considerable variability between different images and different parts of an image in terms of brightness, stain quality and focus quality. We trained the texture classifiers using such data, which made the detection robust to normal variations in image conditions. Thus, the alignment proceeded well, despite the use of classifiers for some landmarks that may have been suboptimal with many false positives (Fig. 3c). The confident detection of the characteristic textures of many structures allowed specimen-specific deviations from the current reference atlas to be discovered and contributed 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 occurred when a landmark was relocated to a nearby region with a similar texture. For example, the right oculomotor nucleus (3N) may have been incorrectly aligned

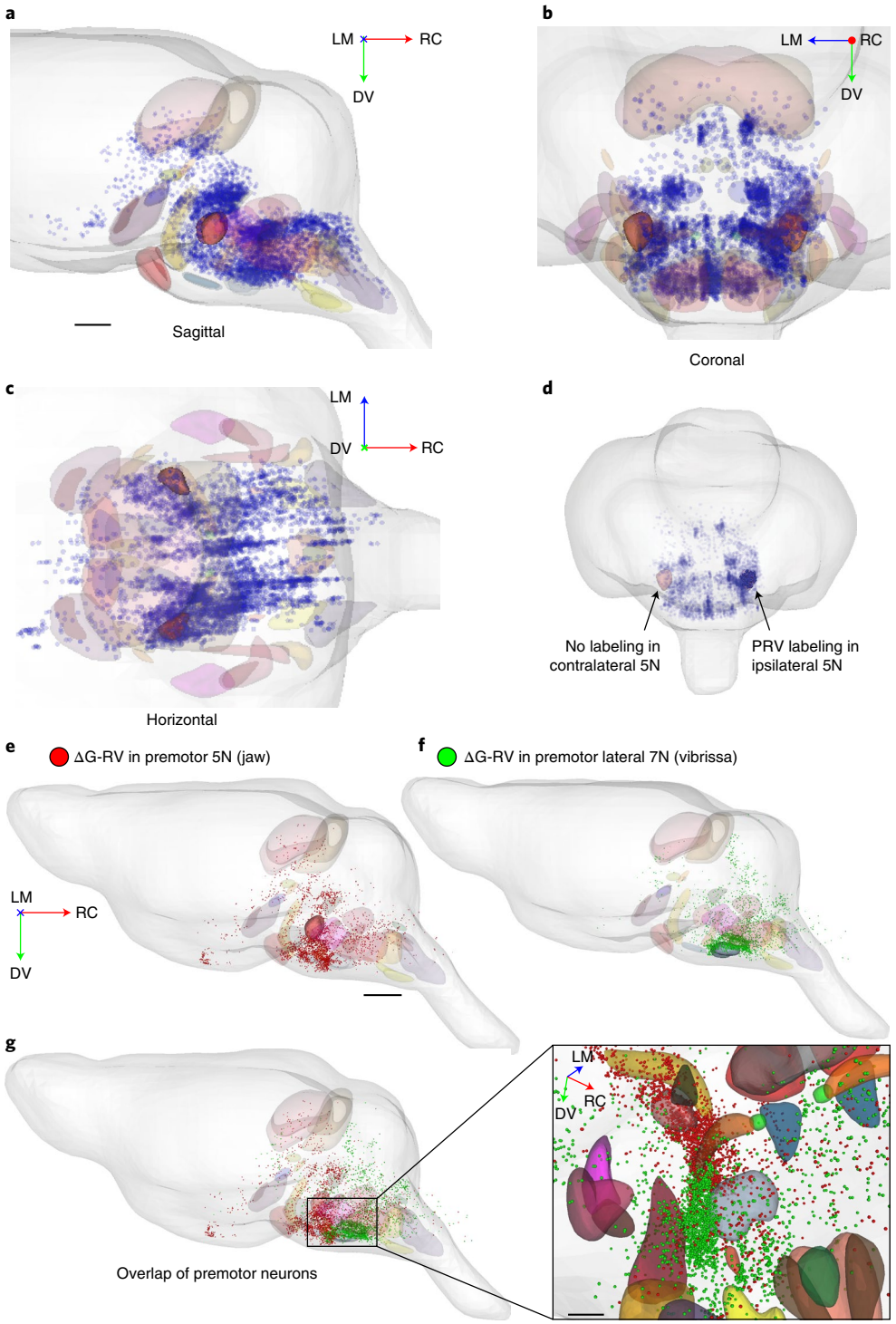

Fig. 5 | Application of the texture-based alignment to fluorescent imaging within and across brains. a–d, Visualization of motor and premotor neurons for a jaw muscle across all three planes in one mouse brain (**a**–**c**). Scale bar, 1 mm. The masseter muscle was injected with PRV that expressed GFP and visualized with a Neurotrace background stain. The PRV labeled cells were manually annotated; alignment with the reference atlas is automatic. **e**–**g**, Visualization of the labeling of different populations of premotor neurons, each in a separate brain, with overlapped density in the parvocellular reticular formation (two pairs of mice). We labeled the premotor neurons of the jaw region of the trigeminal motor nucleus (5N) using G-protein deleted rabies (ΔG-RV) that expressed GFP (red points from one animal; **e**) and the premotor neurons of the vibrissa region of the facial motor nucleus (7N) using the same construct (green points from one animal; **f**). Scale bar, 1 mm in **e**. The inset shows a magnified view of the overlap of the two premotor populations. Scale bar, 500 µm.

with the left ocular motor nucleus immediately adjacent to it. Registration could have failed when a structure was incompletely represented in the images. This could have occurred for structures that were represented in very few sections, such as the abducens motor nucleus (6N), which was as little approximately 50 µm in extent compared with the 20 µm section thickness. Last, registration

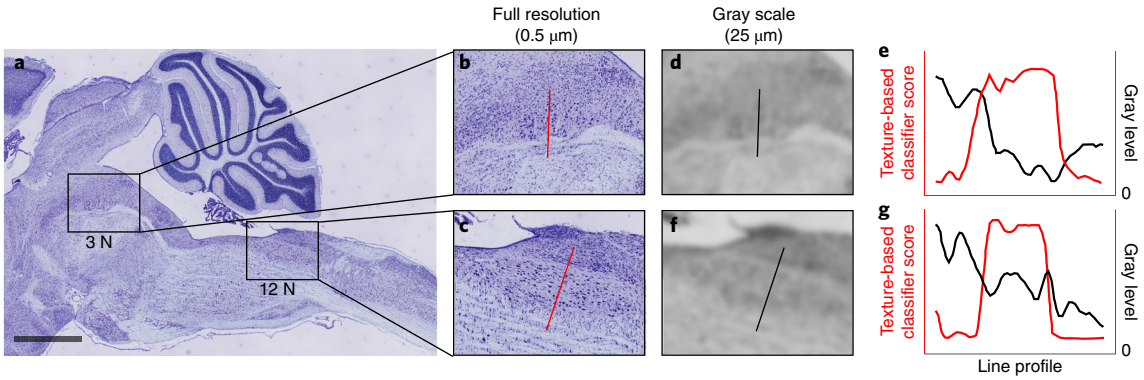

Fig. 6 | Defining a landmark on the basis of texture versus grayscale. The grayscale was formed by downsampling from 0.5 to 25 μm resolution. **a**, Image of the hindbrain and midbrain of a single section stained with thionin. The boxes contain two motor nuclei: the oculomotor nucleus (3N) and the hypoglossal nucleus (12N). Scale bar, 1 mm. **b**, The area around the oculomotor nucleus at higher magnification and a resolution of 0.5 μm. **c**, The area around the hypoglossal nucleus. **d,** The same area as in **b** 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. **e**, Comparison of the probability of the landmark for the oculomotor nucleus from the texture-based classifier versus the value of gray levels along the lines across the nucleus in **b** and **d**. All values are normalized. **f–g**, The analysis of the hypoglossal nucleus performed identically to that of the oculomotor nucleus (**d**,**e**).

could have failed when the textures were diffuse. For example, trigeminal subregions Sp5I and Sp5O were difficult to locate because the boundaries between the subregions were not clearly defined. This inherent difficulty was reflected in their relatively low classification accuracy compared to other structures. In practice, incorrect registration of individual landmarks was rare and was corrected by human verification (Figs. 1 and 3).

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

There were three contributions to the amortization of labor. First, the alignment of new brains with the atlas was automatic except for a verification step. Second, verification involved moving three-dimensional landmarks through the reconstructed volume of all serial sections of a new brain. Last, the verification steps were suitable for accomplishment by less experienced anatomists than those needed for the initial annotation.

Special challenges of the brainstem. The brainstem contains multiple discrete, well-delineated cranial nerve nuclei that served 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. The reticular formations are the site of premotor and pre-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 localized labeled cells and projections on the basis of their triangulation to landmarks that respected the underlying variability from brain to brain (Fig. 5)

Extensions. Our method is applicable to the entire vertebrate brain and to the spinal cord, where the issue of ill-defined boundaries is especially acute. More generally, alignment based on texture can provide the underlying computational engine for mature annotation systems and databases[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 too low in resolution for texture-discrimination[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, clearing and microscopy will always improve the practice of mapping, the current work provides a demonstrated means for automated, high-resolution alignment.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41592-019-0328-8.

Received: 28 June 2018; Accepted: 25 January 2019;
Published online: 11 March 2019

References

1. Roland, P. E. & Zilles, K. Brain atlases: a new research tool. *Trends Neurosci.* **17**, 458–467 (1994).
2. Jones, E. G., Stone, J. M. & Karten, H. J. High-resolution digital brain atlases: a Hubble telescope for the brain. *Ann. N.Y. Acad. Sci.* **1225S1**, E147E159 (2011).
3. MacKenzie-Graham, A. et al. A multimodal, multidimensional atlas of the c57bl/6j mouse brain. *J. Anat.* **204**, 93102 (2004).
4. Majka, P. & Wojcik, D. K. 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).
6. Pauli, W. M., Nil, A. N. & Tyszka, J. M. A high-resolution probabilistic in vivo atlas of human sub-cortical brain nuclei. *Sci. Data* **5**, 180063 EP (2018).

7. Toga, A. W. et al. Postmortem cryosectioning as an anatomic reference for human brain mapping. *Comput. Med. Imaging Graph.* **21**, 131–141 (1997).
8. Swanson, L. W. & Bota, M. Foundational model of structural connectivity in the nervous system with a schema for wiring diagrams, connectome, and basic plan architecture. *Proc. Natl Acad. Sci. USA* **107**, 20610–20617 (2010).
9. Jones, E. G. Viewpoint: the core and matrix of thalamic organization. *Neurosci.* **85**, 331–345 (1998).
10. Braitenberg, V. *On the Texture of Brains, An Introduction to Neuroanatomy for the Cybernetically Minded* (Springer, Heidelberg, 1977).
11. Gong, H. et al. High-throughput dual-colour precision imaging for brain-wide connectome with cytoarchi-tectonic landmarks at the cellular level. *Nat. Commun.* **7**, 12142 (2016).
12. Economo, M. N. et al. A platform for brain-wide imaging and reconstruction of individual neurons. *eLife* **5**, e10566 (2016).
13. Richardson, D. S. & Lichtman, J. W. Clarifying tissue clearing. *Cell* **162**, 246257 (2015).
14. Wilt, B. A. et al. Advances in light microscopy for neuroscience. *Ann. Rev. Neurosci.* **32**, 435–506 (2009).
15. Gray, P. A. Transcriptional factors define the neu-roanatomical organization of the medullary reticular formation. *Front. Neuroanat.* **7**, 1–21 (2013).
16. McElvain, L. E. et al. Circuits in the rodent brainstem that control whisking in concert with other orofacial motor actions. *Neurosci.* **368**, 152–170 (2018).
17. Chiang, A.-S. et al. Three-dimensional reconstruction of brain-wide wiring networks in *Drosophila* at single-cell resolution. *Curr. Biol.* **21**, 1–11 (2011).
18. Peng, H. et al. Brainaligner: 3D registration atlases of *Drosophila* brains. *Nat. Methods* **8**, 493–498 (2011).
19. Ronneberger, O. et al. Vibe-z: a framework for 3D virtual colocalization analysis in zebrafish larval brains. *Nat. Methods* **9**, 735–742 (2012).
20. Randlett, O. et al. Whole-brain activity mapping onto a zebrafish brain atlas. *Nat. Methods* **12**, 1039–1046 (2015).
21. Pinskiy, V. et al. High-throughput method of whole-brain sectioning, using the tape-transfer technique. *PLoS ONE* **10**, e0102363 (2015).
22. Ioffe, S. & Szegedy, C. Batch normalization: accelerating deep network training by reducing internal co-variate shift. *Proceedings of the 32nd International Conference on International Conference on Machine Learning* **37**, 448–456 (2015).
23. Fay, R. A. & Norgren, R. Identification of rat brain-stem multisynaptic connections to the oral motor nuclei using pseudorabies virus. i. Masticatory muscle motor systems. *Brain Res Brain Res. Rev.* **25**, 255–275 (1997).
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 Res.* **585**, 361–366 (1992).
25. Li, Y., Takada, M., Kaneko, T. & Mizuno, N. Premo-tor neurons for trigeminal motor nucleus neurons in-nervating the jaw-closing and jaw-opening muscles: differential differential in the lower brainstem of the rat. *J. Comp. Neurol.* **365**, 563–579 (1995).
26. Mizuno, N. et al. A light and electron microscopic study of premotor neurons for the trigeminal motor nucleus. *J. Comp. Neurol.* **215**, 290–298 (1983).
27. Travers, J. B. & Norgen, R. Afferent projections to the oral motor nuclei in the rat. *J. Comp. Neurol.* **220**, 280–298 (1983).
28. Stanek, E., Rodriguez, E., Zhao, S., Han, B. X. & Wang, F. Supratrigeminal bilaterally projecting neurons maintain basal tone and enable bilateral phasic activation of jaw-closing muscles. *J. Neurosci.* **36**, 7663–7675 (2016).
29. Wickersham, I. R., Finke, S., Conzelmann, K.-K. & Callaway, E. M. Retrograde neuronal tracing with a deletion-mutant rabies virus. *Nat. Methods* **4**, 47–49 (2007).
30. Takatoh, J. et al. New modules are added to vibrissal premotor circuitry with the emergence of exploratory whisking. *Neuron* **77**, 346–360 (2013).
31. Johnson, G. A. et al. Waxholm space: An image-based reference for coordinating mouse brain research. *Neuroimage* **53**, 365–372 (2010).
32. Roland, P. E. et al. Human brain atlas: for high-resolution functional and anatomical mapping. *Hum. Brain Mapp.* **1**, 173184 (1994).
33. Pollack, J. D., Wu, D.-Y. & Satterlee, J. S. Molecular neuroanatomy: a generation of progress. *Trends Neurosci.* **37**, 106–123 (2014).
34. Gonzlez-Vill, S. et al. A review on brain structures segmentation in magnetic resonance imaging. *Artif. Intell. Med.* **73**, 45–69 (2016).
35. Papp, E. A., Leergaard, T. B., Calabrese, E. & Johnson, G. A. Waxholm space atlas of the Sprague Dawley rat brain. *Neuroimage* **97**, 374–386 (2014).
36. MacKenzie-Graham, A. et al. The informatics of a c57bl/6j mouse brain atlas. *Neuroinformatics* **1**, 397–410 (2003).
37. Yushkevich, P. A. et al. Using MRI to build a 3D reference atlas of the mouse brain from histology images. In *Proc. International Society of Magnetic Resonance in Medicine* Vol. 13 (Wiley, 2005).
38. Oh, S. W. 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).
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).
42. Dempsey, B. et al. Mapping and analysis of the con-nectome of sympathetic premotor neurons in the ros-tralventrolateral medulla of the rat using a volumetric brain atlas. *Front. Neural Circ.* **11**, 9 (2017).
43. Senyukova, O. V., Lukin, A. S. & Vetro, D. P. Automated atlas-based segmentation of Nissl-stained mouse brain slices. *Programmi. Comput. Soft.* **37**, 245–251 (2011).
44. Amunts, K. & Zilles, K. Architectonic mapping of the human brain beyond Brodmann. *Neuron* **88**, 1086–1107 (2015).
45. Frth, D. et al. An interactive framework for whole-brain maps at cellular resolution. *Nat. Neurosci.* **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 ofthe adult mouse brain. *Nat. Neurosci.* **12**, 356–362 (2009).
49. Mazziotta, J. et al. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philos. Trans. Royal Soc. B* **356**, 1293–1322 (2001).
50. Parekh, R. & Ascoli, G. A. Neuronal morphology goes digital: a research hub for cellular and system neuroscience. *Neuron* **77**, 1017–1038 (2013).
51. Miller, M. I., Beg, M. F., Ceritoglu, C. & Stark, C. Increasing the power of functional maps of the medial temporal lobe by using large deformation diffeo-morphic metric mapping. *Proc. Natl Acad. Sci. USA* **102**, 9685–9690 (2005).
52. Tsai, P. S. et al. Correlations of neuronal and microvascular densities in cortex revealed by direct counting and colocalization of nuclei and vessels. *J. Neurosci.* **29**, 12455314570 (2009).
53. Ragan, T. et al. Serial two-photon tomography for au-tomated ex vivo mouse brain imaging. *Nat. Methods* **9**, 255–258 (2012).
54. Ren, J., Choi, H., Chung, K. & Bouma, B. E. Label-free volumetric optical imaging of intact murine brains. *Sci. Rep.* **7**, 46306 (2017).
55. Tsai, P. S. et al. All-optical histology using ultrashort laser pulses. *Neuron* **39**, 27–41 (2003).

Acknowledgements

The idea for this project was catalyzed at the 2008 meeting on 'The Architectural Logic of Mammalian CNS' at the Banbury Center, Cold Spring Harbor Laboratory. We thank N.M. Lindsay for help with annotating the trigeminus, A. Brzozowska-Prechtl and H. Liechty for assistance with the histology, A. Newberry for assistance with coding, and X. Ji and K. Svoboda for timely discussions. This work was funded by NIH BRAIN awards (U01 grant nos. MH105971 and NS0905905 and U19 grant nos. MH114821 and NS107466), a Mathers Charitable Foundation award, and funds from the Dr. George Feher Experimental Biophysics Endowed Chair. We thank L. Enquist and the Center for Neuroanatomy with Neurotropic Viruses (NIH grant no. OD010996) for supplying the PRV.

Author contributions

Y.F., H.J.K. and D.K. conceived the project. Y.C. and Y.F. designed the algorithm. D.F., B.F., D.K., L.E.M., P.P.M. and A.S.T. contributed data. Y.C., Y.F., B.F., H.J.K., D.K. and L.E.M. planned experiments and analyzed data. Y.C., Y.F., B.F. and D.K. wrote the manuscript. D.K. and P.P.M. dealt with the many institutional organizations that govern animal health and welfare, surgical procedures and laboratory health and safety issues that include specific oversight of chemicals, controlled substances, cutting tools and viruses.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41592-019-0328-8.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to Y.F. or D.K.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

See the Nature Research Reporting Summary for additional details.

Subjects and sample preparation. The dataset for building the atlas consisted of 12 brains of postnatal day 56 (P56) male C57BL/6J mice in which all sections were stained with thionin (Supplementary 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 PRV at a titer of 1×10^9 particles per ml with Neurotrace blue staining and three for additional tests. Last, we used two brains of male ChAT-cre mice (JAX no. 006410) crossed with the FLEX-TVA mice (JAX no. 024708), both sacrificed at P56, with injections of EnvApseudotyped glycoprotein-deleted rabies-eGFP at a titer of 3×10^7 particles per ml (Salk Institute for Biological Studies Virus Core), and two brains of male P56 ChATcre crossed with the Ai14 reporter (JAX no. 007914), both also sacrificed at P56. 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 fixed and 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 per pixel resolution and a digitization depth of 8 bits, or a Zeiss AxioScan Z.1 at 0.325 µm per pixel resolution and a digitization depth of 16 bits. For animals injected with PRV only, the expression of GFP was enhanced 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 only used the portion of the images that contained the brainstem, that is, 270 sections cropped to 20,000 by 15,000 pixels. Last, 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 had minimal large-scale distortion. To align all sections, we first downsampled the images by a factor of 32; to a pixel size of 16 µm. We aligned the sections by computing two-dimensional rigid transforms between every pair of adjacent sections using Elastix[56] with the mutual information as the optimization criteria[57]. The correlation was 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 (Supplementary Fig. 2, top right). 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 to use more flexible deformable transforms. Finally, we used the transform matrices derived from downsampled images to compute transform matrices that corresponded 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 Fig. 2). Manual boundaries were represented by closed polygons and their vertices were recorded. We manually annotated every sections. On average, it took an annotator 1 min to draw one boundary and 60 h to annotate a full brain with the 51 selected landmarks.

Bootstrapping the reference atlas. We converted each set of annotated images of brain sections to a set of three-dimensional binary maps that provided the locations of different pre-averaged landmarks; that is, the landmark from an individual annotated brain. The voxel size of the map was 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 were inside versus outside the pre-averaged 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. 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.

Estimating nominal shapes. To estimate the nominal shape of a landmark, we aligned all instances of the pre-averaged 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 contained each voxel (Supplementary Fig. 4c–f). Intuitively, the reference atlas was defined by situating the centroid of each shape at its corresponding nominal position.

Training texture classifiers. Patches of grayscale, full-resolution images served as inputs to the classifiers. We found that a size around 100 µm, or 224×224 pixels, showed both local brain organization and detailed cell shape. Larger patches were also effective (Supplementary Fig. 12) but may have failed to capture small structures. Patches were collected on the basis of a moving window with a pitch of 32 µm that yielded roughly 40,000 patches per section. Training patches for a certain structure were collected from all sections on which this structure was annotated. A patch was labeled positive if at least three of the four corners were located inside a boundary of this structure (Fig. 2e). Similarly, a negative patch needed to have three corners in the bordering zone of a boundary. The use of negative patches in the boundary region, rather than anywhere in the image, improved 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) to encode the patches. This CNN had previously been trained on a subset of ImageNet, a dataset of 21,000 natural scene images in 1,000 categories, and achieved state-of-the-art classification performance. We modified the network to accommodate single-channel input and used the 1,024-dimensional vector that fed 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 were used. We denote the feature vector of the ith patch by $\mathbf{x}_i$ and its label by y_i (Fig. 2e). The L2-penalized logistic regression minimizes:

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

The optimal weight vector $\mathbf{w}$ 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 $\mathbf{x}$ and the weight vector of a particular classifier is $\mathbf{w}$, 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 a voxel size of 16 µm on the edge (Fig. 3c). The resolution of these volumes was low so that they could 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 (Fig. 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 $\mathbf{A}\in\mathbb{R}^{3\times3}$ and a shift vector $\mathbf{b}\in\mathbb{R}^3$. The transform maps a coordinate $\mathbf{x}$ in the reference atlas to another coordinate $\mathbf{Ax}+\mathbf{b}$ in the input brain.

Denote Φ to be 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:

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

The optimal $\mathbf{A}$ and $\mathbf{b}$ are found by stochastic gradient ascent. At each iteration, the Jacobian is computed based on 10,000 randomly sampled voxels from each structure. The adaptive gradient algorithm Adagrad was employed to automatically control the learning rate. Convergence was usually achieved in fewer than 100 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 computed a rigid transform separately for each landmark. The three-dimensional rigid transform for structure r is denoted by $G(\mathbf{x};\boldsymbol{\omega}_r,\mathbf{u}_r)=\mathbf{R}(\boldsymbol{\omega}_r)\mathbf{x}+\mathbf{u}_r$, where $\mathbf{u}_r\in\mathbb{R}^3$ is the shift vector and $\mathbf{R}(\boldsymbol{\omega}_r)\in\mathbb{R}^{3\times3}$ is a rotation matrix parametrized by the Euler vector $\boldsymbol{\omega}_r\in\mathbb{R}^3$.

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:

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

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.

Timing. In the current version of our system, verification by an astute user took ≃5 min across the 51 landmarks in our current brainstem atlas. Typically, five of the landmarks required a correction, which took ≃1 min per landmark, or about 10 min total after all verification steps. This time should be compared to the roughly 10 min per landmark in the initial annotations or nearly 10 h per initial brain. While human verification of the automatic alignment was the rate-limiting step, the net throughput became 60 times greater compared to the initial annotation of a brain (Supplementary Table 2).

Deformation field. 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 (Supplementary Fig. 11), in and outside all landmarks. For location $\mathbf{x}$, the deformation vector is expressed as:

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

where $\mathbf{c}_r$ is the centroid of landmark r after alignment, $Z(\mathbf{x}) = \sum_r b(\|\mathbf{x}-\mathbf{c}_r\|)$, 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 (Fig. 3).

In manually annotated brains, the landmark structures derived from automatic alignment were compared to manual annotations, using the isosurface for a probability of $P = 0.05$. 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-dimensional 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 (equations 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° around three axes.

Peak width. We computed the Hessian matrix of the objective function at the maximum with respect to translations in three directions. On the basis of the eigenvalues and eigenvectors of the Hessian, we derived the most certain and the least certain translation directions that 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 chose moderately sized windows that were evenly spaced across image. For each window we computed a linear correction factor to make the pixel values have a zero mean and unit standard deviation. Correction factors across adjacent windows were interpolated to make the correction smooth over each image. In detail, 2 by 2 mm windows were taken across an image with 1.2 mm even spacing. For each window the mean, μ, and standard deviation, σ, of the pixel values were 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 used the reference atlas to reduce annotation time. Using the in-house program (Supplementary Fig. 2), 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 two-dimensional structure boundaries were 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 were used for training and testing.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Code availability

All analysis was done following the algorithms detailed in the Methods. The code was written in Python and is available as Supplementary Material and, together with updates, at https://github.com/ActiveBrainAtlas/MouseBrainAtlas through the GNU General Public License (GPL). Organization of the code is in a ReadMe file.

Data availability

All raw data are publicly available. They may be downloaded, with a listing of files found in Supplementary 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.

References

56. Klein, S., Staring, M., Murphy, K., Viergever, M. A. & Pluim, J. P. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans. Med. Imag.* **29**, 196–205 (2010).
57. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G. & Suetens, P. Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imag.* **16**, 187198 (1997).

nature research

Corresponding author(s): David Kleinfeld and Yoav Freund

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Data was acquired with two commercial slide scanning systems. (1) Nanozoomer-XR. We use the Nanozoomer proprietary software for data acquisition and the Kakadu SDK library (kakadusoftware.com) to transform the raw NDPI format data into loseless JPEG2000 files. (2) Zeiss Axio Scan.Z1. We use Zen proprietary software for data acquisition and the freely available CZItoTIFFConverter program (cifweb.unil.ch/index.php?option=com_content&task=view&id=152&Itemid=2) to transform the raw data to TIFF.
Data analysis	All analysis was done following the algorithms described in detail in the Methods section. The code is written in Python and is available at https://github.com/ActiveBrainAtlas/MouseBrainAtlas.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All raw data is publicly available. Currently, it may be downloaded from https://neurophysics.ucsd.edu/ActiveBrainAtlas/, with a listing of files found in Table 1 and at https://github.com/ActiveBrainAtlas/MouseBrainAtlas/blob/master/doc/Brain_stack_directories.md. Upon acceptance of the manuscript, data will be made available through the Amazon Web Service Storage S3 at the bucket named mousebrainatlas-rawdata.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	26 mouse brains. This led to a determination that systematic errors surpassed statistical errors
Data exclusions	Five brains were lost as a result of failure of the staining. This is a pre-established criteria.
Replication	We successfully tested against 12 replicates after the initial atlas was formed
Randomization	Not relevant here.
Blinding	All annotators worked in isolation and were blinded to group allocation

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	For the PRV amplification, we used Novos Biologicals NB600-303 anti-GFP primary (https://www.novusbio.com/products/gfp-antibody_nb600-303) as detailed in text.
Validation	Validation was in terms of labeling of mouse motoneurons in the trigeminal motor nucleus (5N).

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	The dataset for building the atlas consists of 15 brains of postnatal day 56 (P56) male C57BL/6J mice in which all sections were stained with thionin. We used an additional 11 brains of P56 male C57BL/6J mice (JAX no. 000664), five solely for alternate sections of thionin and Neurotrace blue staining and four brains of male ChAT-cre mice (JAX no. 006410) crossed with the FLEX-TVA mice (JAX no. 024708), all sacrificed at P56, with injections of a fluorescent viral agent, and three brains of male P56 ChAT-cre crossed with the Ai14 reporter (JAX no. 007914).
Wild animals	*Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*
Field-collected samples	*For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*