Received: 29 April 2022 | Revised: 23 August 2022 | Accepted: 1 October 2022

DOI: 10.1002/hbm.26126

TECHNICAL REPORT

WILEY

An artificial-intelligence-based age-specific template construction framework for brain structural analysis using magnetic resonance images

Dongdong Gu[1] | Feng Shi[1] | Rui Hua[1] | Ying Wei[1] | Yufei Li[2,3] | Jiayu Zhu[4] | Weijun Zhang[4] | Han Zhang[5,6] | Qing Yang[5,6] | Peiyu Huang[7] | Yi Jiang[6] | Bin Bo[2] | Yao Li[2] | Yaoyu Zhang[2] | Minming Zhang[7] | Jinsong Wu[8,9] | Hongcheng Shi[10] | Siwei Liu[10] | Qiang He[4,11] | Qiang Zhang[4] | Xu Zhang[6] | Hongjiang Wei[2] | Guocai Liu[12] | Zhong Xue[1] | Dinggang Shen[1,5,13] | the Consortium of Chinese Brain Molecular and Functional Mapping (CBMFM)

Correspondence
Zhong Xue and Dinggang Shen, Shanghai United Imaging Intelligence Co., Ltd., 701 Yunjin Road, Xuhui District, Shanghai, China.
Email: zhong.xue@ieee.org and
Email: dgshen@shanghaitech.edu.cn

Funding information
Shanghai Zhangjiang National Innovation Demonstration Zone Special Funds for Major Projects "Human Brain Research Imaging Equipment Development and Demonstration Application Platform", Grant/Award Number: ZJ2018-ZD-012; National Natural Science Foundation of China, Grant/Award Numbers: 61901256, 82027808, 62071176, 91949120

Abstract

It is an essential task to construct brain templates and analyze their anatomical structures in neurological and cognitive science. Generally, templates constructed from magnetic resonance imaging (MRI) of a group of subjects can provide a standard reference space for analyzing the structural and functional characteristics of the group. With recent development of artificial intelligence (AI) techniques, it is desirable to explore AI registration methods for quantifying age-specific brain variations and tendencies across different ages. In this article, we present an AI-based age-specific template construction (called ASTC) framework for longitudinal structural brain analysis using T1-weighted MRIs of 646 subjects from 18 to 82 years old collected from four medical centers. Altogether, 13 longitudinal templates were constructed at a 5-year age interval using ASTC, and tissue segmentation and substructure parcellation were performed for analysis across different age groups. The results indicated consistent changes in brain structures along with aging and demonstrated the capability of ASTC for longitudinal neuroimaging study.

KEYWORDS
artificial intelligence, brain template, longitudinal template, magnetic resonance imaging (MRI), registration, statistical analysis

For affiliation refer to page 873

1 | INTRODUCTION

In the past 30 years, magnetic resonance imaging (MRI) techniques have enabled imaging of human brains in high spatial resolution, aiming at accurately measuring the shapes and sizes of the brain and its substructures, studying structural and functional connectivity, understanding human cerebral development and degeneration, and assisting treatments of brain diseases such as brain tumor, stroke, and neurodegenerative disorders. Brain image registration is often used to automatically navigate and quantify intricate brain structures and facilitate group analysis, as well as to perform neurological disease evaluation, diagnosis, and prognosis. The construction of structural templates from different groups of images is also essential for understanding the brain anatomy in neurology and cognitive science.

By registering a group of images onto a common image space and forming a template, one can establish anatomical correspondences across individuals, allowing for analyzing and comparing images in a group or between different groups. Therefore, for longitudinal studies, constructing age-specific MRI templates is beneficial to neurological research by capturing the trend of anatomical changes during the development and degeneration courses in a population. The creation of age-specific brain templates is more complicated as not only the brains change with aging but also anatomical structures vary across different subjects. More generally, age-specific templates may help decouple individual variability with aging characteristics, and subsequent analysis such as brain tissue segmentation and substructure parcellation based on age-specific templates could better quantify age-related diseases from healthy ones.

In the literature, early attempts to construct brain templates were based on annotating the volumetric image of the individual subject, often represented in terms of size, shape, and image intensity of the human brain (Brodmann, 1909; Talairach, 1988). The Talairach atlas was the commonly used human brain template developed from sections of a 60-year-old French female. It created a standardized grid for neurosurgery and can be employed to identify deep brain structures in stereotaxic coordinates. Today, the Talairach atlas is still commonly used in neuroimaging, but the lack of a three-dimensional model makes it difficult to map locations from three-dimensional anatomical MRI to the atlas automatically.

To better analyze and understand the similarity and variation of brain structures within a population, image registration is often applied to generate a template by warping all the images onto a common space, which can be fixed or iteratively updated. The Montreal Neurological Institute (MNI) atlases were constructed based on several hundreds of subjects using registration techniques (Evans et al., 1993; Mandal et al., 2012; Mazziotta et al., 2001). MNI305 (Evans et al., 1993) is the average of linearly registered 305 normal brain MRIs using 9 degree of freedom (DoF) affine transformations and is considered the first MNI atlas. In 1998, a lab member of MNI (Colin Holmes) scanned his brain 27 times and linearly registered them to MNI305 to create an average atlas (Colin27) (Holmes et al., 1998). Colin27 has since been used as the standard template in SPM (Friston, 2003). Later, to generate the MNI linear template (9 DoF), also known as the International Consortium for Brain Mapping template, 150 high resolution and contrast-improved MRI volumes of young adults were linearly registered to the MNI305 space. In 2009, MNI152 was constructed with 152 images using nonlinear registration (Mazziotta et al., 2001). Atlases of T1, T2, PD, and brain tissue maps are available for MNI152. The construction procedure involved 40 rounds of iterations, wherein individual volumes are nonlinearly registered to the average of the warped images obtained from the previous iteration. Similarly, templates were proposed for the cohorts of infants and children in Shi et al. (2014), Shi et al. (2011), shi et al. (2010), and Zhang et al. (2016).

Accurate image registration plays an important role in template construction, but the ones built upon a single reference image could introduce a systematic bias toward the shape of that image. Group-wise registration (Wu et al. 2011), therefore, takes the role to solve a population center, while aligning all the images onto the common space (Wu et al., 2012), which results in more accurate and consistent results for better investigating group similarity and variation.

One of the most popular group-wise registration pipelines is first calculating the average of the warped images and then registering all subject volumes to this average image, and iteratively refining the template by repeating the registration and averaging procedure. However, the results may be undermined by the blurry average image at the beginning as all the subject volumes are far from being well-aligned, which might lead to slow convergence in the downstream registration due to the lack of anatomical details. For better initialization, Park et al. (2005) selected a template closest to the geometrical mean among a population using multidimensional scaling and registered other subject images onto the template. Another method is to first perform pair-wise image registration and then solve a group average of deformations as the initialization for the group-wise registration (Seghers et al., 2004). The former might be still biased to the selected subject, while the latter could bring a computational burden, especially for large population size. Additionally, simultaneous registration of all the subject images based on a global cost function optimization can be performed (Learned-Miller, 2005; Q. Wang et al., 2010; Zitova & Flusser, 2003), but the high dimensionality of variables makes the optimizer easy to be stuck into local minima.

To overcome large anatomical differences between images, methods have been proposed to register each subject with the group center image with the help of intermediate templates, which are more similar to the subject. In light of this, Kim et al. (2011) and Baloch and Davatzikos (2009) and Baloch et al. (2007) generated intermediate templates and combined the deformation from the template to the intermediate template with that from the intermediate template to the subject. The method was also applied to diffusion tensor images (DTI) in Jia et al. (2011). Similarly, Munsell et al. (2009) built a minimum spanning tree (Kruskal, 1956), wherein each node represents an image, and each edge corresponds to the distance between two nodes or the image similarity such as intensity mean square error, to help bridge the registration between the template to each subject.

With the network constructed among subject images, manifolds formed by the image set provide intrinsic relationships among the

subjects, from which the pseudo-geodesic median image can be selected as the template (Hamm et al., 2009). Because any neighboring image pair in the network are similar to each other, large anatomical differences between the template and a subject can be well-aligned through a network path connecting them along the network edges. Based on the learned manifold, ABSORB (Jia et al., 2010) was proposed to warp each subject image to the template step by step, and the population center can also be updated iteratively. In this way, each subject only deforms locally to its qualified neighbors according to a learned manifold, and large shape differences can be better tolerated during the registration.

Another task for template construction is to estimate a sequence of templates to capture the trend of anatomical changes in a population, which is particularly valuable for studying neural development and degeneration from the fetus, infant, or children, to the elderly (Dittrich et al., 2014; Fonov et al., 2011; Habas et al., 2010; Kuklisova-Murgasova et al., 2011; Serag et al., 2012; Zhang et al., 2016). In 2010, a Chinese brain atlas Chinese_56 (Tang et al., 2010) was created based on 56 male volunteers from 20 to 30 years old. Later, to increase the number of samples and to cover different ages, the Chinese2020 atlas (Liang et al., 2015) was constructed with over 2000 subjects from multiple centers. The templates were constructed in 12 different age groups using T1-weighted images. Recently, G. Yang et al. (2020) constructed age-, sex-, and sample size-matched Caucasian and Chinese adult brain/head volumetric templates with probability tissue maps of gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) to quantify the impacts of sample sizes and population on brain template construction, and to assess the morphological differences between Caucasian and Chinese.

Image segmentation plays an important role in template construction. The commonly used brain parcellation methods can be categorized into three types: multi-atlas registration-based parcellation (Aljabar et al., 2009; Artaechevarria et al., 2009; Klein & Hirsch, 2005; Shi, Fan, et al., 2010), machine learning-based parcellation using handcrafted features (Ashburner & Friston, 2005; Pereira et al., 2016; van Opbroek et al., 2013; L. Wang et al., 2015), and deep learning (DL)-based methods (Chen et al., 2018; Çiçek et al., 2016; He et al., 2016; Stollenga et al., 2015). The multi-atlas-based methods can align multiple templates to the target image and fuse the respective warped label maps to yield the final labels. Registration is often time-consuming and may suffer from registration errors. Machine learning methods employ classifiers such as the support vector machine (van Opbroek et al., 2013), Gaussian mixture models (Ashburner & Friston, 2005), or random forests (L. Wang et al., 2015) with spatial and intensity features. However, the handcrafted features may suffer from limited representation capability for accurate recognition considering the large structural variation of the brain. Data-driven DL-based methods can learn the feature representation, and networks such as 3D convolutional neural network (CNN) (Çiçek et al., 2016), long short-term memory (Stollenga et al., 2015), and deep residual learning (Chen et al., 2018; He et al., 2016) prompted the brain parcellation in an end-to-end manner. These DL methods can achieve accurate segmentation results and do not require any manually designed features.

No matter what template construction strategies, image registration is inevitably the building block of the pipeline. Earlier research only used linear registration, which could fail to accurately map the anatomies to the template. Conventional deformable registration such as Demons (Vercauteren et al., 2009), HAMMER (Shen & Davatzikos, 2002; Zacharaki et al., 2008), ORBIT (Zacharaki et al., 2008) and SyN (Avants et al., 2008) were later used in template construction. However, iterative optimization procedures made the computing slow. DL-based registration has been investigated to improve computational efficiency while providing comparable accuracy (Fu et al., 2020; Gu et al., 2020; Litjens et al., 2017). To train registration networks, supervised learning is based on carefully building/simulating the ground-truth deformations (X. Cao et al., 2018; X. Cao et al., 2017; Eppenhof & Pluim, 2018; Lv et al., 2018; Rohé et al., 2017; Sokooti et al., 2017; Uzunova et al., 2017; X. Yang, 2017), and unsupervised learning (Balakrishnan et al., 2018, 2019; de Vos et al., 2019; Ferrante et al., 2018; Hu et al., 2019; Sedghi et al., 2018; Sokooti et al., 2017; Stergios et al., 2018) is more popular as it does not require ground-truth deformations but learns by maximizing the similarity between images, such as the sum of squared difference and cross-correlation (CC), together with deformation regularization losses. The inferring of DL registration often takes only a few seconds. Therefore, we integrated DL registration and segmentation in this work.

An artificial-intelligence-based age-specific template construction (ASTC) framework is presented to generate templates for different age groups using MRIs in this article. T1, T2, and T2 FLAIR images of 646 subjects from 18 to 82 years old were collected from four medical centers. The procedure of ASTC consists of three major steps: (1) age-specific brain templates were constructed using DL-based registration with a 5-year age interval; (2) state-of-the-art DL-based brain parcellation algorithm was applied for image segmentation; and (3) because the templates were average warped images, a DL-based image mapping network trained from a separate image data set was presented to sharpen image contrast and to provide clear anatomical structures.

After computing the T1-based templates, we statistically analyzed the volumes of brain regions and their longitudinal changes. The results show that volumes of anatomical structures obtained from the templates are consistent with the measures obtained from the original images and also show consistent aging trends. We also applied the deformation fields to generate T2 and T2 FLAIR template images as well as the tissue segmentation maps (in the supplemental material).

This article is organized as follows. Section 2 presents the detailed steps of ASTC, including iterative registration for template construction and image sharpening using CNN. In Section 3, we describe the experiments to evaluate the constructed templates. Sections 4 and 5 are the discussion and conclusion of this study, respectively.

2 | METHODS

2.1 | The framework for ASTC

Age-specific brain templates are beneficial to analyze brain development and degeneration disorders. Generally, the templates of

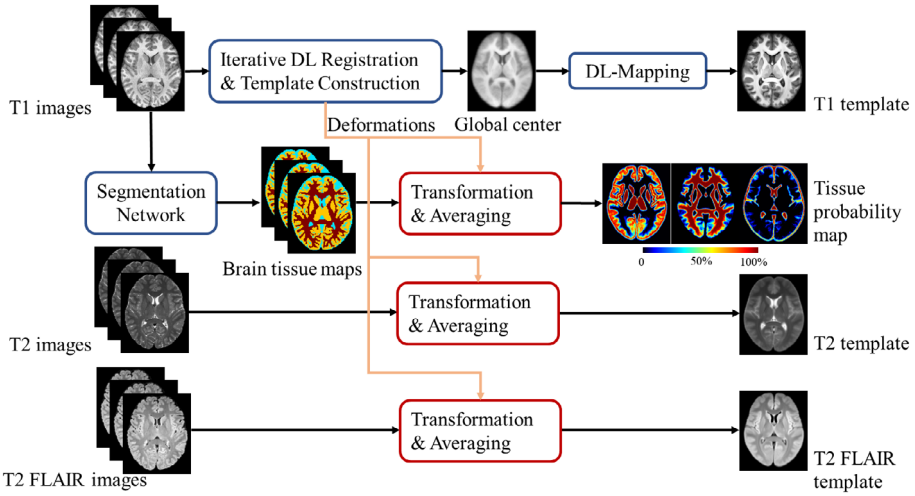


FIGURE 1 The artificial-intelligence-based age-specific template construction (ASTC) pipeline for one age group. First, iterative deep learning (DL)-registration is used to compute the T1 template, followed by DL-mapping for image enhancement (T1 enhanced template); then, the segmentation network is applied for tissue segmentation; and T2, T2 FLAIR templates and brain tissue probability maps are obtained by using the deformation fields from T1 registration

different ages should not only reflect common anatomical structures in each age group but also consistently capture the longitudinal changes. Therefore, after grouping the subjects according to ages, our objective is to generate a template for each age group by defining the voxel-wise correspondences between the template and each subject, which should be accurate enough to reflect a clear subcortex and gyral structures in the population center.

Figure 1 shows the ASTC pipeline, where the sample images are from the 25-year group. The template for each age group was constructed independently, and totally 13 templates from 20 to 80 years old at a 5-year interval were constructed. Given a group of T1 images, first, the template was constructed using an iterative DL-registration algorithm. As the averaged images are often blurry in appearance, we implemented an image sharpening procedure to enhance the T1 template using a DL-mapping network. The deformation between the T1 template and each subject was recorded for transforming the corresponding T2 and T2 FLAIR images onto the template space. The T1 images were also segmented into WM, GM, and CSF maps using a CNN segmentation network and transformed onto the template space. In this way, we obtained the T1, T2, and T2 FLAIR templates; the T1 enhanced template, as well as the tissue probability maps for the age group. The enhancement of T2 and T2 FLAIR templates could be performed in the same way as the T1 mapping, but the procedure was not implemented in this article.

The detailed registration framework, segmentation network, and sharpening methods will be discussed below.

2.2 | ASTC algorithm details

The images were grouped according to age, and 13 templates were constructed from 20 to 80 years old at a 5-year interval. For each group, intersubject registration was first performed to bring all the subject images roughly to a common space, that is, the global center. In preprocessing, we first aligned all the T1 images onto the MNI152 atlas using 9 DoF linear transformations and B-Spline interpolation. We refer to these globally aligned images as the original subject images.

The iterative registration procedure for an age group is shown in Figure 2. Basically, we used the iterative registration and group mean method (Joshi et al., 2004) to update the average warped image to obtain the final template. A DL-based group-wise image registration network (Gu et al., 2020) was employed to perform intersubject registration, wherein the conventional losses like the smoothness and the mean squared image intensity similarity losses, as well as the inverse-consistency loss were used for training the network. In the first iteration (Figure 2a), the initial template (blue) was set as the average of the inputs (green), and DL-registration was performed to register the input images onto the template, resulting in respective warped images (yellow). Because all the structures of the brain may not be perfectly aligned, the warped images are not exactly the same as the template. However, compared to the original subject images, they are much similar to the template. Therefore, the new average of the warped images (yellow) was used as the new template in the next iteration. In the second iteration (Figure 2b), we applied DL-registration to register the new template with each subject image (i.e., the warped images from iteration 1 represented by yellow) and obtained the new warped images (red). The new template can thus be updated again. In this way, the template was iteratively updated. In this work, we applied two iterations to obtain the final template as the warped images have a similar appearance. The final deformation between the template and each subject can be the composition of the deformation fields obtained in each iteration.

For T2 and T2 FLAIR images and brain tissue maps, we first aligned them to T1 images using 6 DoF rigid registration. To generate the final templates, global and deformable transformations were combined into one by composing the rigid transformation matrix with the deformation fields, and image transformation and interpolation were only performed once to minimize partial volume effects. Finally, the templates of T2 and T2 FLAIR images were obtained by averaging the warped images, and the brain tissue probability maps were obtained in the same way.

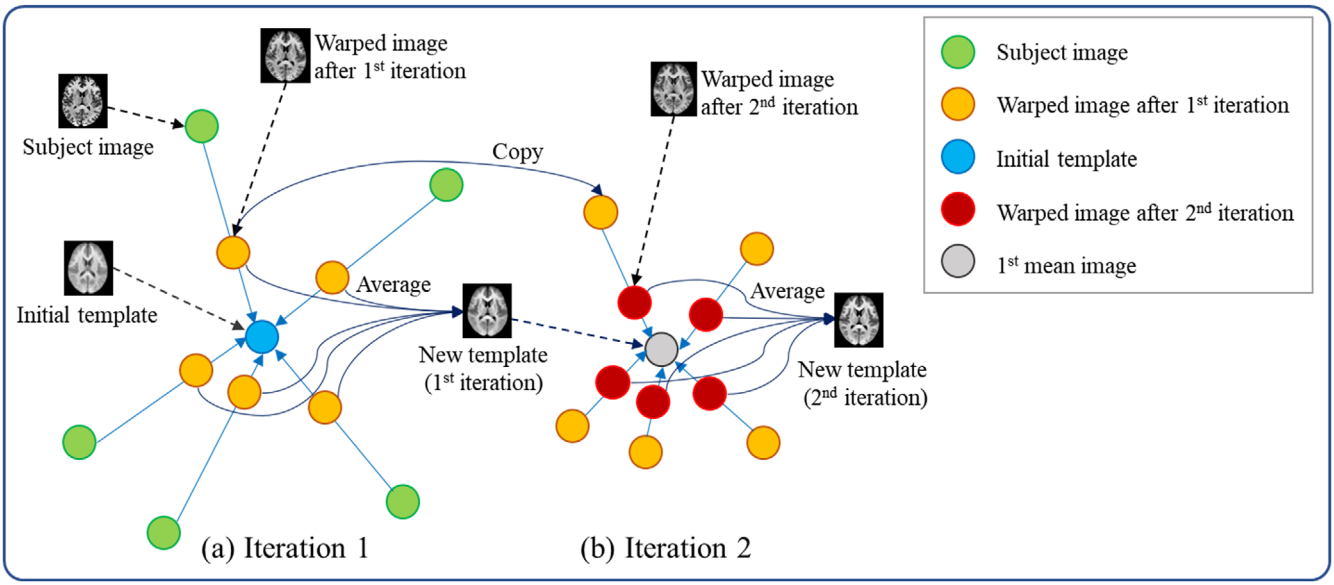
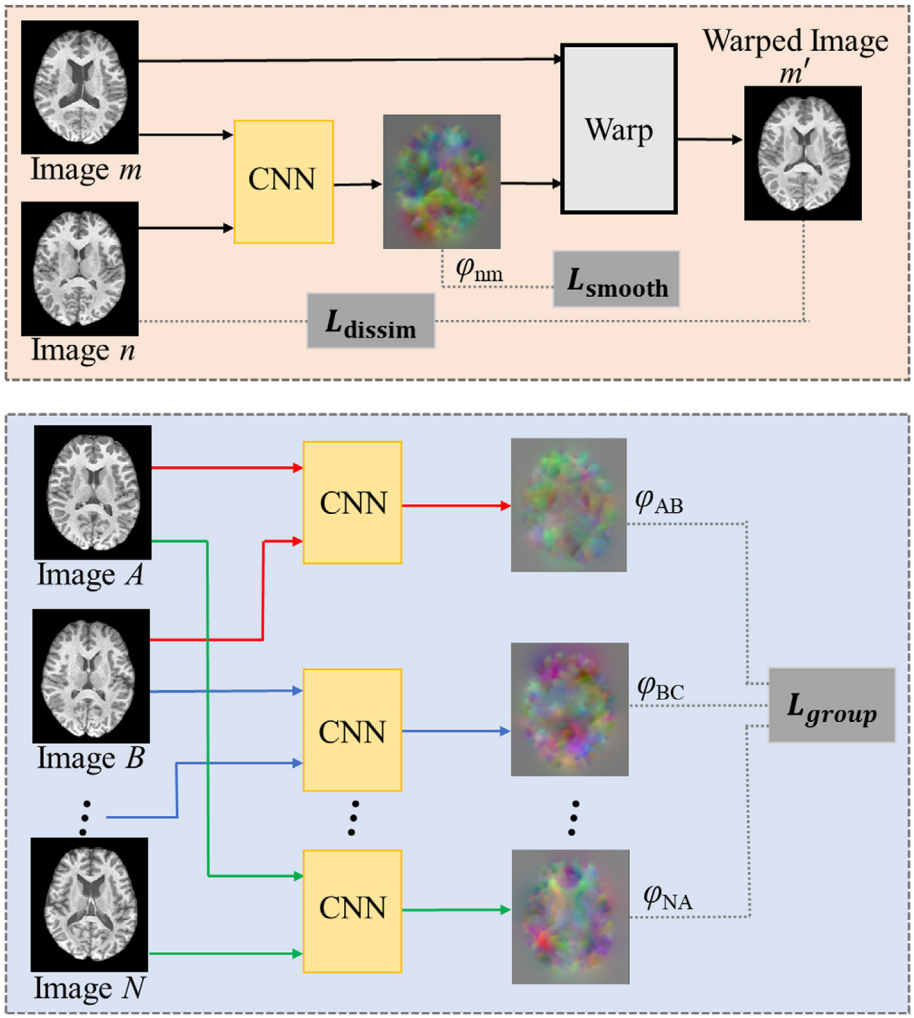


FIGURE 2 Iterative registration in artificial-intelligence-based age-specific template construction (ASTC). (a) In the first iteration, the average image (blue) of all the input images (green) is calculated and registered with the input images, and the warped images are represented in yellow, whose average is used as the new template in (b) (gray); (b) all the images (yellow) are registered onto the new template (gray), which provides the new warped images (red) to be averaged as another new template. In this way, the final template can be generated from the input image group.

FIGURE 3 Group-wise consistent registration network. Top: pair-wise training; and bottom: cycle consistency training



2.2.1 | Consistent registration network

For intersubject deformable registration, we trained a group-wise consistent CNN to preserve cycle consistency in the deformations among the images (Gu et al., 2020). The training strategy of this network is shown in Figure 3. The CNN structure is similar to U-net with a pair of images *m* and *n* as the input, and the output is a deformation field φ_{nm}. The "warp" block was implemented using the spatial transform network (STN) (Jaderberg et al., 2015), which implemented trilinear deformation of the moving image, so that the loss function can be back-propagated for network training.

There were two steps during the training. The first applied the regular unsupervised registration loss, that is, the dissimilarity loss L_{dissim}, between the warped subject image n' and the template image *m*. The loss can be measured by mean square error or CC. We also used the weighted sum of the deformation gradient and Jacobian determinate as the smoothness loss L_{smooth}. However, as shown in Gu et al. (2020), the network trained may not yield consistent deformation fields if one swap the order of the two input images. Therefore, we applied a group-wise consistency loss in the second step. Specifically, for a properly ordered image group $(A, B, \ldots, N)$, the adjacent two images can be aligned sequentially to get a series of deformations, for example, $\varphi_{\text{AB}}, \varphi_{\text{BC}}, \ldots, \varphi_{\text{N}-1\rightarrow\text{N}}$. Group consistency means that, when composing all the deformations, the resultant field would be equal or similar to the deformation estimated by directly registering the N^{th} and the first images. Therefore, a cycle-consistency constraint was employed as an additional loss L_{group} based on the closed deformation loop, and the composition of deformations, $\varphi_{AB}\circ\ \varphi_{BC}\ \ldots\ \varphi_{N-1\rightarrow N}\circ\ \varphi_{NA}$, should be close to an identity field. In this work, we set *N* to 3 to form a loop for any given three images in the group. The registration can reduce systematic bias caused by the order of input images, increase robustness, and improve the reliability of data analysis by incorporating this constraint.

The initial network parameters were obtained from Gu et al. (2020) trained using 100 images from a separate data set, and the network was trained further based on the above-mentioned training strategy using the images included in this study. Once the registration network is well-trained, we applied the models and calculated the average warped image for each age group (as shown in Figure 2). The registration can be accomplished in 10 s for each image pair with size $182 \times 218 \times 182$ to greatly accelerate the template construction procedure.

2.2.2 | Brain tissue segmentation and brain parcellation network

For brain tissue segmentation and brain parcellation, we segmented the preprocessed T1 images into GM, WM, and CSF, as well as 106 regions of interest (ROIs) by using a pretrained cascade VB-Net, which was designed for medical image segmentation using a typical CNN with a U-net structure (Wei et al., 2021; Xiao et al., 2019). All the T1 images were preprocessed by skull stripping, bias correction, and resampling to $1 \times 1 \times 1\ \text{mm}^3$. Images were manually delineated by neurologists for evaluation. The segmentation network combined coarse localization and segmentation refinement and was successfully used in medical image segmentation tasks including the brain tumor (Hua et al., 2020) and thoracic organs (Han et al., 2019). The most time spent for brain tissue segmentation is reading and writing images, which is about 1 s for one input. The segmentation network itself costs only about 500 ms. On average, the segmentation can be finished in 1.6 s.

2.2.3 | Template enhancement network

As mentioned above, we incorporated group-wise consistency in registration training and iteratively updated the average image to generate the template, whose quality increased iteratively. However, the appearance of such an averaged image is still blurry (as shown in Figure 5) because any registration error or subtle anatomical structure changes could contribute to unclear anatomy structure in the template. Thus, we designed an image enhancement network in order to enhance the image quality.

Specifically, we trained a CNN, which contains ResBlock modules for image mapping as in Lan et al. (2020). We modified the network structure to 3D with image patches of size [32, 32, 32] adaptive to GPU memory. The size of input and output images is the same, as the goal is to enhance image quality instead of super resolution. We applied an anisotropic convolution strategy, given that the main focus of the task was to improve the image fidelity between slices. We used two consecutive convolution kernels with shapes $1 \times 3 \times 3$ and $3 \times 1 \times 1$ instead of the traditional $3 \times 3 \times 3$ kernel during training. The network included 32 ResBlock modules, and the number of channels is 256. The loss function was defined as the weighted sum of mean squared error and the structural similarity index measure between the output and the target images (Liu et al., 2021).

The training samples for image enhancement were generated as follows. Since the deformation fields from the template and subject images were known, we can deform all the subject images to any subject space. As shown in Figure S1, the deformation field from image I_1 to I_2 is the composition of the inverse of field φ_1 and field φ_2. The deformation field from I_1 to I_3, ... I_n can be composed in the same manner. After deforming all the images to subject I_1, the warped images are in the same space with I_1 and can be averaged to generate the blurry image $I_{\text{f\_1}}$ in the space of image I_1. Thus, for subject *t*, $I_{\text{f\_t}}$ can be formulated as:

$$I_{\text{f\_t}} = \frac{1}{N-1}\sum_{x=1}^{N(x\neq t)} I_x\left(\varphi_x \circ \varphi_t^{-1}\right), \tag{1}$$

where φ_t is the deformation field from the template to image I_t, and φ_x denotes the deformation field from the template to subject *x*.

In this way, we can generate image pairs for each subject (a blurry image and the original subject image), and the enhancement network can be further trained using these samples. In order to have more

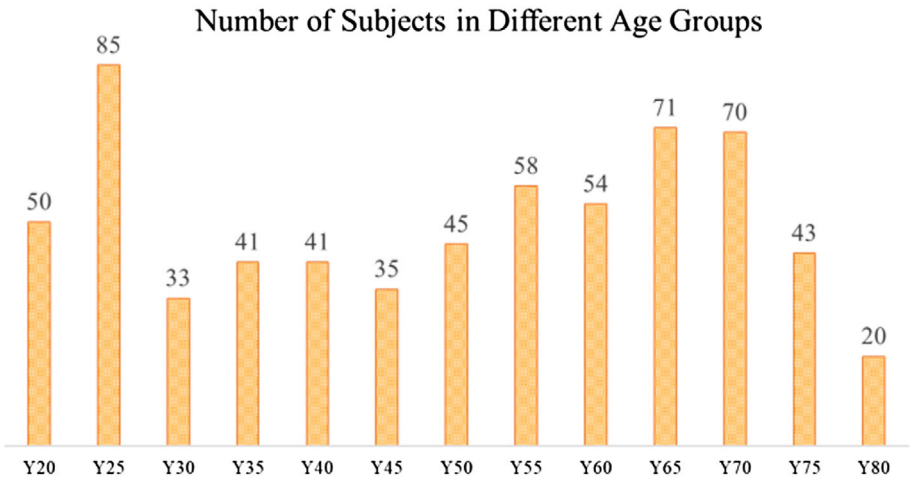


FIGURE 4 Age distribution of the dataset

training samples, only one enhancement network was trained by combining all the sample images constructed using this method. Finally, the template image for each age group was enhanced by using the average template as the input. The enhancement procedure can be finished in 1.75 min and is performed only once for generating each age group's template. The experiments were performed on Ubuntu 16.04.1 LTS, with 1 NVIDIA TITAN Xp 12 G GPU and 32 Intel (R) Xeon(R) CPU E5-2620 v4 @ 2.10 GHz using PyTorch 1.8.1 and CUDA 10.0.

2.3 | Data and preprocessing details

The data set used for constructing the templates was collected from the ongoing Chinese brain molecular and functional mapping (CBMFM) project, aiming to develop the brain templates of Asians in the aging process. In this article, we used T1, T2, and T2 FLAIR images of 646 healthy subjects (334 female and 312 male participants with an age range of 18–82 years). None of the participants reported a history of psychological or neurological disorders, and they can autonomously respond to neuropsychological scale tests. This article focused on the longitudinal changes in brain structure with age, and gender differences were not taken into account. According to the approved Institutional Review Board protocol (No. 2019-105-01), all the volunteers signed informed consent forms for participating in this study. All the T1-weighted MRI data were acquired using 3.0 T UIH uMR790 scanner at four medical institutions, as the Institute of Brain-Intelligence Technology at the Zhang Jiang Laboratory, Institute for Medical Imaging Technology of Shanghai Jiaotong University, Huashan Hospital of Fudan University, and the Second Affiliated Hospital of Zhejiang University.

To evaluate the interscanner reproducibility, 10 traveling subjects were scanned using the standardized CBMFM protocols at the four sites involved in this project. The consistency in scanners' performance is measured using intraclass correlation coefficients (ICC) (Fisher, 1992) and coefficients of variance (CV). For cortical thickness assessment, the mean and standard deviations (stds) of ICC across all the cortical ROIs are 0.79 and 0.14, respectively. In surface area estimation, the mean of the regional averaged CV values is 2.69% and the std is 1.36%. For regional volumes, the whole-brain averaged ICC is 0.91 (std = 0.09). The results showed good consistency and low variation in brain volume measurements. T1 images from 263 subjects were quality controlled (QC) using the method proposed in (Esteban et al., 2019). The QC metrics from the CBMFM project with those from other large-scale data sets (e.g., the autism brain imaging data exchange (Nielsen et al., 2013) and the consortium for reliability and reproducibility (Zuo et al., 2014)) were compared, and the results demonstrated the stability and reproducibility of the MRI scanners.

The detailed scanning parameters were as follows: repetition time = 8.07 ms, echo time = 3.4 ms, inversion time = 1060 ms, acquisition matrix = 320 × 300, flip angle = 8°, number of slices = 208, spacing = 0.8 × 0.8 × 0.8 mm^3. T2-weighted images and T2 FLAIR images were also acquired using the same scanner with consistent parameters as spacing of 0.8 × 0.8 × 0.8 mm^3 for T2 images and 1.0 × 1.0 × 0.7 mm^3 for T2 FLAIR images, respectively. The age distribution is shown in Figure 4.

The trained model of the DL-based group-wise image registration network was the same as in Gu et al. (2020), where 100 images used for training were different from our longitudinal data set, these images formed 4950 different image pairs for training. The model's parameters were then used as initialization for further refining the registration for each age group. It is worth noting that the objective of template construction is to come up with a common-space image representing the group in an unsupervised way, so registration does not need to be evaluated in a traditional training–testing strategy.

For registration, the network was trained using 4950 image pairs from a separate data set in an unsupervised manner. We applied the registration model and calculated the average warped image for each age group in the template construction procedure. Specifically, we first trained the group-wise consistent deep registration network using the 4950 image pairs as shown in Figure 3. During testing, the initial template for each age group was the average of input images in the first iteration. The average of the warped images was used as the new template in the next iteration. Then, we applied the DL-

registration to register the new template with each subject image (i.e., the warped images from iteration 1) and obtained the new warped images. The new template can thus be updated again. The template was iteratively updated. The final deformation between the template and each subject is the composition of the deformation fields obtained in each iteration. The template of each age group was constructed separately based on this procedure.

For segmentation, a total of 1889 images with delineation from experts were used for training and 295 cases were used for testing. These T1 images were collected from our collaborating institutes as a separate data set particularly used for segmentation. It turned out that the mean Dice coefficient of the segmented brain structures with the ground-truth labels (delineation from experts) was 0.857, and the standard deviation was 0.051 for the testing data set. The mean values of GM, WM, and CSF were 95.0, 96.9, and 92.7, respectively. After the model is trained and validated using separate data, we applied it to segment brain tissues in our age-specific MRI scans, that is, the samples in 13 different age groups. The 646 subject images herein are regarded as the testing samples to apply forward propagation using the well-trained segmentation network to obtain the brain tissue segment results.

As for preprocessing, we applied N4 bias correction, skull stripping, intensity normalization (to the range of 0–255), histogram matching, and affine alignment with the MNI152 atlas for T1 images. T2-weighted and FLAIR images were aligned to their corresponding T1 images using 6 DoF rigid registration with B-Spline interpolation. Notice that rigid registration might not exactly match different modalities, and deformable registration could be applied in the future to quantitatively evaluate the age-group templates for these mapped modalities. Skull stripping was performed on T1 images by training a deep segmentation network using FSL BET (Jenkinson et al., 2012) results as the ground truth. Brain regions in T2 and FLAIR images were extracted using the brain masks of the T1 images. All the final preprocessed images were with an intensity range between 0 and 255 and isometric spacing $0.8 \times 0.8 \times 0.8\ mm^3$.

Basically, compared to traditional registration, by packaging a neural network and STN using GPU, the output deformation field can be obtained in about 10 s for two input images with size $182 \times 218 \times 182$. The image warping itself can be finished in less than 1 s (interpolation is the only calculation). Traditional optimization-based registration methods could take about 60 s (Demons) and 40 min (SyN) to obtain the deformation field in our experiments for the same image pairs.

For image enhancement, the initial model was first trained from 642 subjects, the same as in our previous image super-resolution reconstruction paper (Z. Cao et al., 2021) using different samples. Then, the model was further refined to enhance the templates in this study. By generating a blurry image for each subject as shown in Figure S1 and Equation (1), together with the corresponding original subject image, we obtained image pairs to serve as training/refining samples. Only one enhancement network was trained by using all the sample images to guarantee large number of training samples. Finally, after the model was well trained, 13 template images for the 13 age groups were enhanced. Notice that in this case, the 13 averaged templates are not used for training and are regarded as the testing samples to go through the enhancement network and provide sharpened template images with detailed cortical information.

For cortical thickness and subcortical volume extraction, we used the "recon-all -s *subject_file_path* -all -qcache" command of FreeSurfer (7.2.0 Release Version for Ubuntu 18), the computational time for each subject required several hours.

3 | RESULTS

We first used ASTC to generate T1 templates. Figure 5 shows the axial view of the 30-year group template after affine registration, after the first and the second iterations, as well as after image enhancement. It can be seen that the appearance of the template obtained using affine registration is very blurry especially for the cortical areas, as it is obtained by directly averaging the affine transformed images.

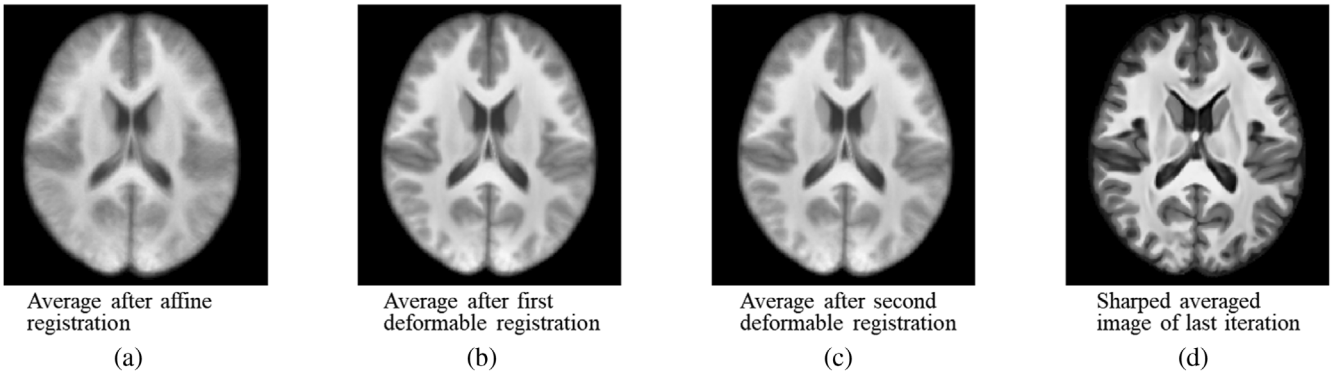


FIGURE 5 An example of the evolution of the average images during the course of template construction for the 30-year group. (a) Initial template by averaging all the affined transformed images, (b) the template after the first round of deep learning (DL) registration, (c) the template after iteration 2, and (d) the template after DL mapping

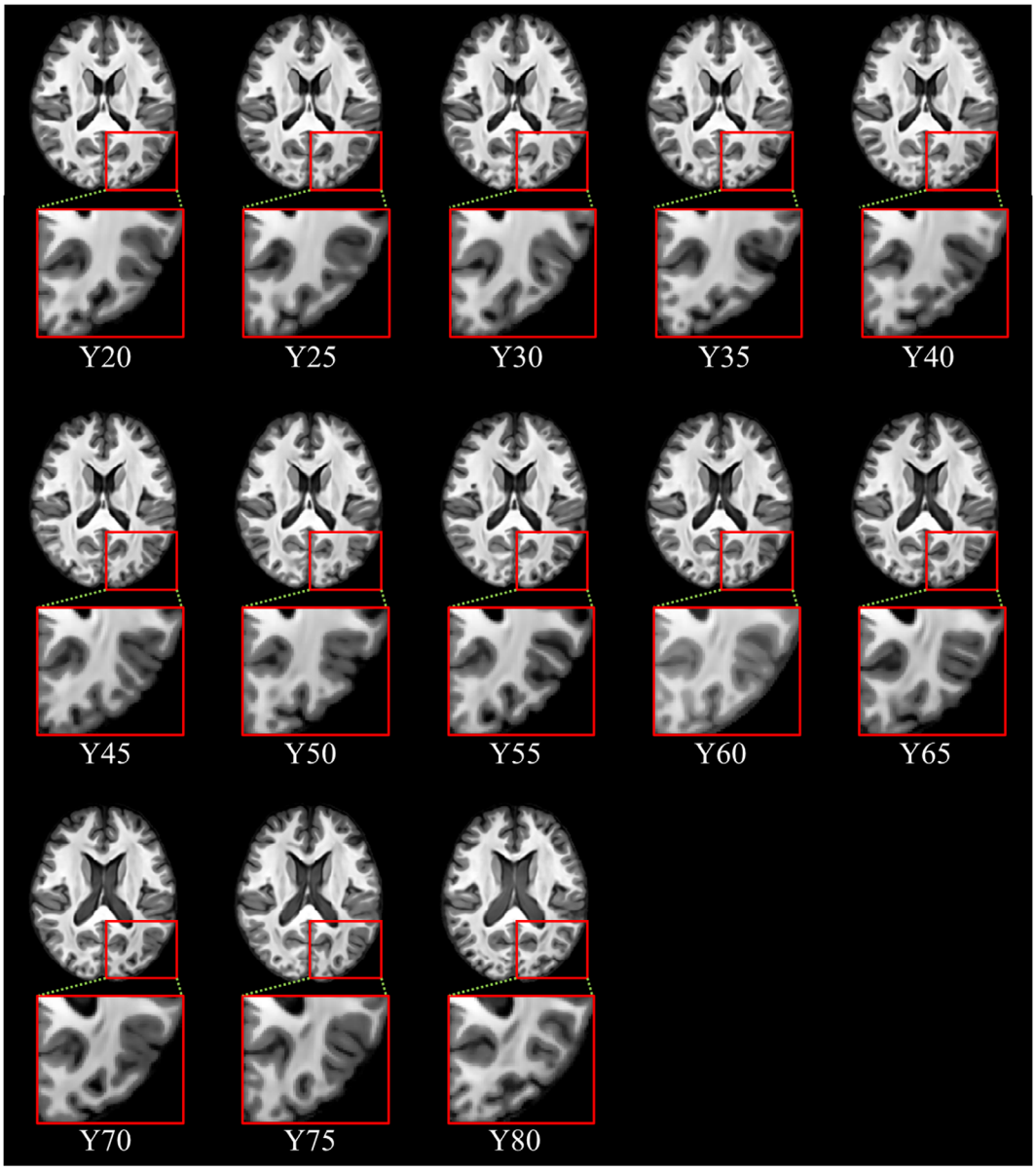


FIGURE 6 Sharpened T1 weighted longitudinal template. The red boxes mark zoomed in details

The template can be iteratively refined by using DL image registration and becomes clearer after each iteration. However, the result is still not satisfactory along the cortical surface.

Taking the advantages of image mapping technology, the DL-based image sharpening network can learn the appearance and cortical structures from a large number of samples and make the averaged warped images clear enough to identify brain structures, as shown in Figure 5d. The 3D rendering volumes of the averaged images during the course of template construction are provided in the supplementary material (Figure S2). It can be seen that the cortical structures become clear with the iteration and sharpening procedure. The reason for different gyral patterns across different age groups could be that images used for different age groups are from different subjects. Although iterative reconstruction generates a template that is close to the samples, such templates could be biased to the brain shapes of the subjects within their group, especially when the number of samples is small.

Using the same ASTC procedure, we generated 13 T1 templates for different age groups from 20 to 80 years old at an age interval of 5. For T2-weighted templates, T2 FLAIR templates and tissue distribution maps, the deformation fields generated for T1 templates were applied to get the warped maps, and the averages of their warped versions were recorded as the templates. Herein, because the major objective in this article was to analyze T1 templates, DL mapping was not used for T2 and T2 FLAIR images.

We generated sharpened T1 templates over the whole age range, as shown in Figure 6. T1 templates, T2-weighted templates and T2 FLAIR templates are provided in the supplementary material (Figure S3, Figure S4, and Figure S5), and only the axial views are

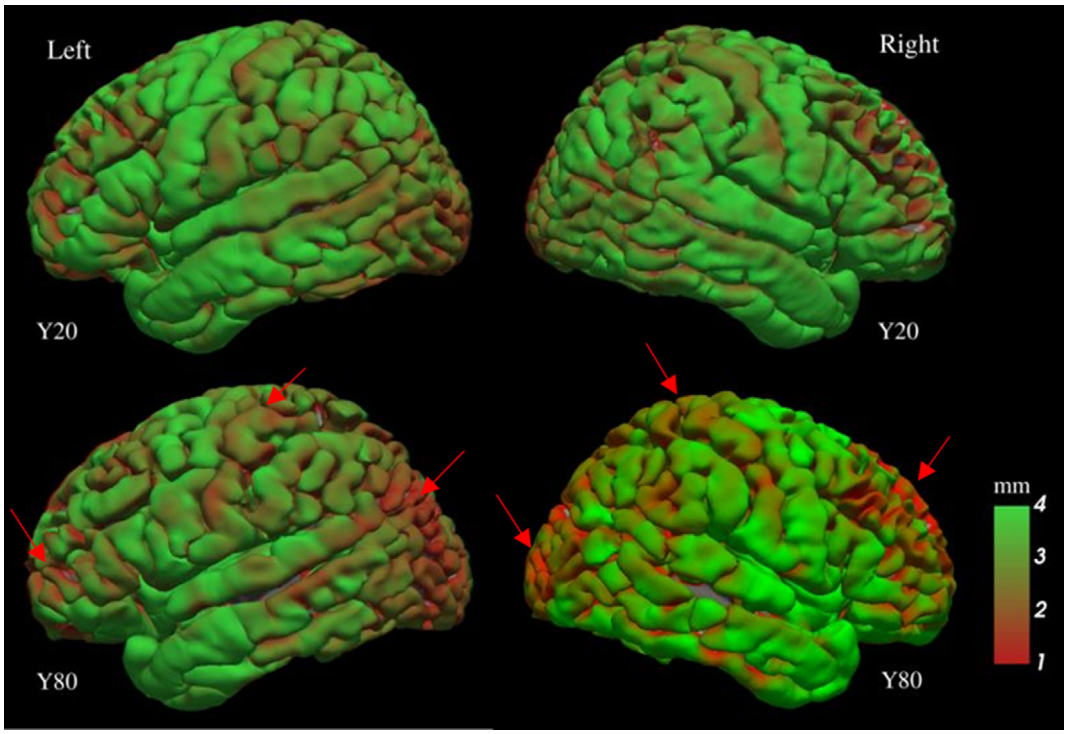


FIGURE 7 Cortical thickness maps of 20- and 80-year-old template. Lateral views of left and right hemispheres are shown, respectively

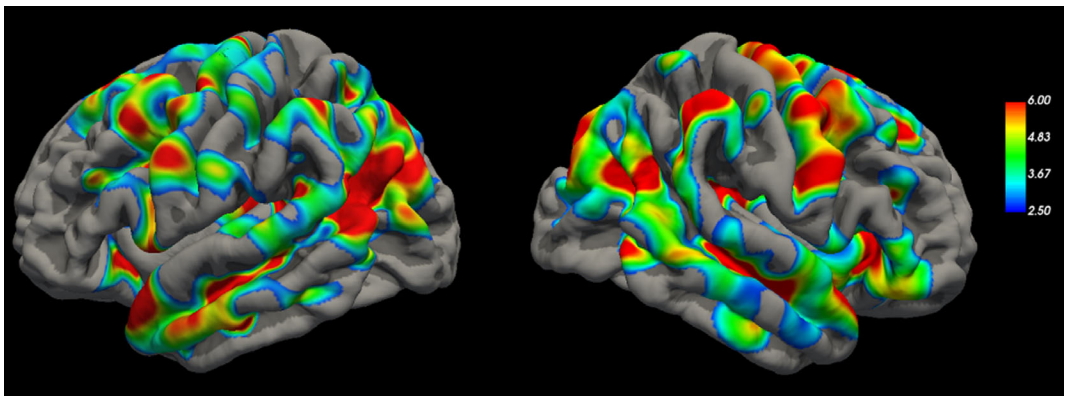


FIGURE 8 Regions with significant cortical thickness differences between fifty 20- and twenty 80-year-old subjects

illustrated. In each slice, the template images show clear anatomical structures in the cerebral tissue, subcortical regions, cerebellum, and brainstem. Similar morphological structures can be found in the brain gyri and sulci across the longitudinal templates, while slight anatomical differences are present in the peripheral GM and WM regions (Figure S6). This is because the images used for constructing the age-group templates are from different subjects, and the number of subjects is still considerably small. It can also be seen that the size of the ventricle enlarges with age, especially after 60 years old. WM atrophy can also be visible from the images.

Cortical thickness maps of Y20 and Y80 are shown in Figure 7. FreeSurfer (Fischl, 2012) was used for constructing and computing the thickness using the sharpened T1 weighted longitudinal templates. An age-related decrease in cortical thickness can be observed in the figure. The most decreased cortical thickness locations are in the frontal lobe, occipital lobe, and paracentral lobules (indicated by the red arrows). The results are consistent with the demonstrations in Frangou et al. (2022) and Podgórski et al. (2021).

Additionally, one-sided unpaired statistical test was performed followed by cluster-correction with vertex-wise threshold $-\lg(p) > 3$, and cluster-wise p-value threshold, $p < .05$. The regions with significant differences are shown in Figure 8. The colormap shows voxel-wise p-value in $-\lg(p)$ scale. The results reveal three clusters on the left hemisphere: superior temporal, superior frontal, and

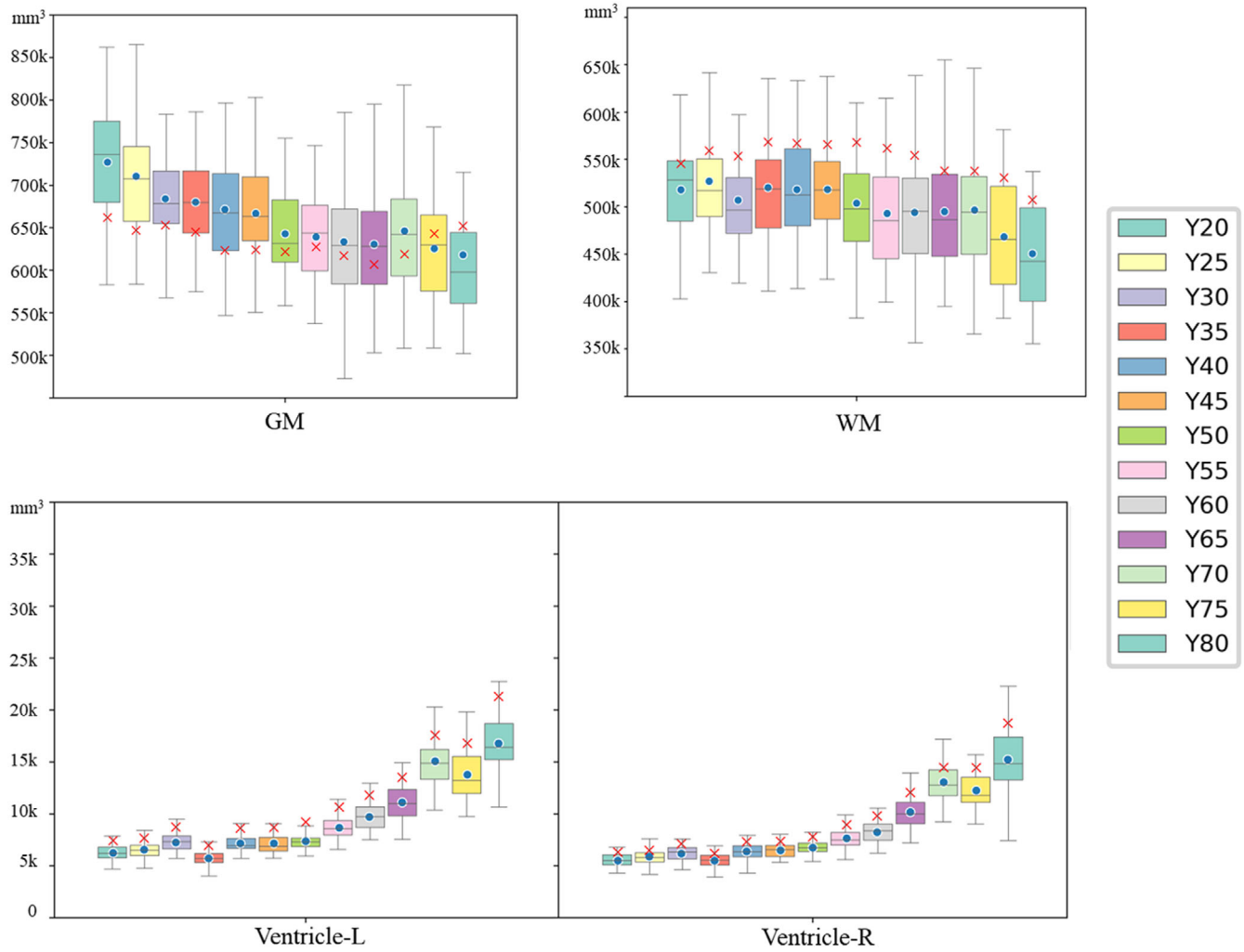


FIGURE 9 Boxplots of volumes of gray matter (GM)/white matter (WM)/ventricle across all age groups

parahippocampal gyrus, and four clusters on the right hemisphere: superior temporal, fusiform, rostral middle frontal, and superior parietal gyri.

For tissue volumes, we calculated the volumes of WM, GM, and the lateral ventricles for each subject and for the templates before and after sharpening. Figure 9 shows the boxplots of volumes of different tissues types of the subjects. It can be seen that the volumes of GM decrease with age, and those of WM decrease after 45-year-old group. The ventricles show an enlarged tendency with age. The red points "x" and blue points "○" inside the boxes are volumes computed from the templates before and after template sharpening. It can be seen that the blue points "○" are close to the median values of the boxplots, reflecting the longitudinal volumes better reflect the population than the red points "x." This could be due to discrepancies caused by segmenting blurry templates. Thus, image sharpening could potentially yield better tissue contrast and provide better volumetric measures.

Similarly, Figure 10 shows the boxplots of volumes of hippocampus, parahippocampal gyrus, amygdala, caudate, thalamus, putamen, and pallidum in both hemispheres of the brains. The overall trends of volumes of these structures are downward. These trends are similar with the results in Dima et al. (2022), Hedman et al. (2012), and Scahill et al. (2003). The blue points inside the boxes are the mean values of each brain tissue or structural volume in the age-group templates. We normalized the volumes to the subject images space according to the affine matrixes to eliminate the individual differences. It can be seen that the blue points are almost in the middle of each box, which means that our constructed templates can reflect the longitudinal volume distribution of the human brain.

The brain templates are freely available (https://github.com/Duoduo-Qian/longitudinal_brain_atlas).

4 | DISCUSSION

It is worth noting that other algorithms such as variational autoencoder (VAE) (Kingma & Welling, 2019) and generative adversarial network (GAN) (Antipov et al., 2017) can be used to generate more images and brain structures based on available samples. However, given relatively small number of samples in each age group, learning

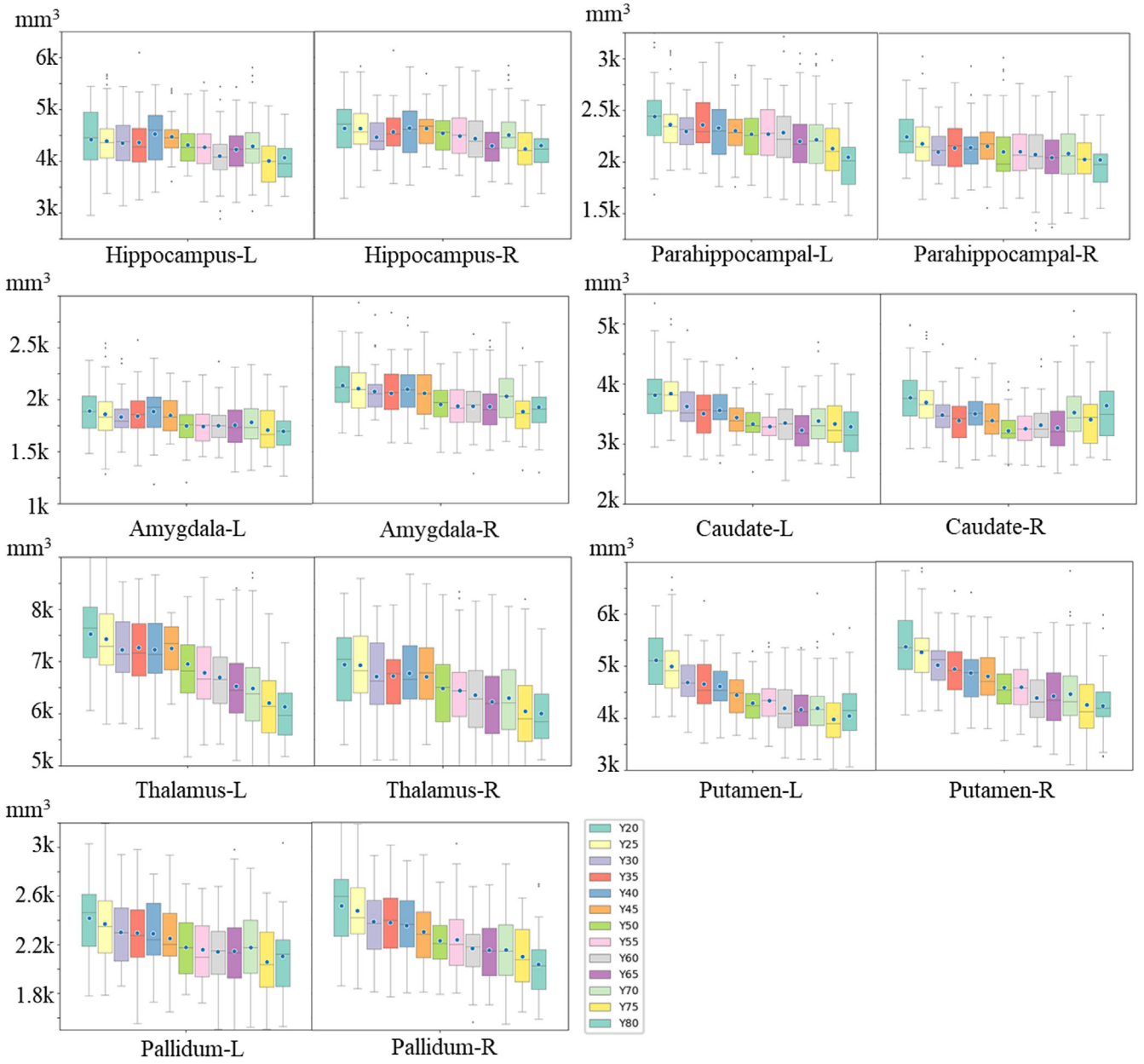


FIGURE 10 Boxplots of volumes of brain structures across all age groups

and analyzing the age-specific shape variability could be challenging and could lead to the loss of subtle information. Therefore, we chose to directly register the images and compute the template using iterative DL warping. With the superior performance of VAE and GAN for large number of samples, we believe that they could be used for estimating the variability of age-related brain images in future work with more samples be collected.

For the image enhancement, we used ResBlock modules to learn the residuals between different layers. In the loss function, the structural similarity measure between the output and the target images is designed in addition to the mean squared error, which can capture detailed information about the complex brain structures (Z. Cao et al., 2021). In this case, although we tried GAN, its training was unstable in our experiments, and structural shapes seem not very stable for us (comparing the input and output images).

5 | CONCLUSION

This article presented an ASTC framework for brain structural analysis using MRIs. In ASTC, DL-based iterative group-wise consistent registration network and image sharpening network were used to generate morphological plausible longitudinal templates. Segmentation and parcellation, and tissue probability maps were also generated together with volumes of anatomical structures for auxiliary quantitative analysis. Altogether, 13 sets of multimodality template images were

computed from 20 to 82 years old at a 5-year interval. The results showed that the volumes of anatomical structures obtained from the templates were consistent with those measured from the original images, indicating that they can capture major structural characteristics of different age groups. The trend of volumetric changes was also consistent with the aging studies of normal brains. In future studies, we plan to incorporate longitudinal constraints for template construction instead of individual age groups and explore other age-related templates (e.g., DTI and functional MRI) and take the disease conditions into account for longitudinal analysis.

AFFILIATIONS

[1]Shanghai United Imaging Intelligence Co., Ltd., Shanghai, China
[2]School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China
[3]School of Mathematics and Computer Science, Chifeng University, Chifeng, China
[4]Shanghai United Imaging Healthcare Co., Ltd., Shanghai, China
[5]School of Biomedical Engineering, ShanghaiTech University, Shanghai, China
[6]Institute of Brain-Intelligence Technology, Zhangjiang Lab, Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai Center of Brain-Intelligence Engineering, Shanghai, China
[7]Department of Radiology, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China
[8]Glioma Surgery Division, Neurologic Surgery Department, Huashan Hospital, Shanghai, China
[9]Medical College, Fudan University, Shanghai, China
[10]Department of Nuclear Medicine, Zhongshan Hospital, Fudan University, Shanghai, China
[11]United Imaging Research Institute of Innovative Medical Equipment, Shenzhen, China
[12]Hunan University, Changsha, China
[13]Shanghai Clinical Research and Trial Center, Shanghai, China

ACKNOWLEDGMENTS

This work was supported by Shanghai Zhangjiang National Innovation Demonstration Zone Special Funds for Major Projects "Human Brain Research Imaging Equipment Development and Demonstration Application Platform" (ZJ2018-ZD-012), and partly by the National Natural Science Foundation of China (61901256, 82027808, 62071176 and 91949120).

DATA AVAILABILITY STATEMENT

The study was part of the ongoing Chinese Brain Molecular and Functional Mapping (CBMFM) project, aiming to develop brain templates in the aging process. According to the approved IRB protocol (No. 2019-105-01), all the volunteers signed informed consent forms for participating in this study.

ORCID

Dongdong Gu https://orcid.org/0000-0002-0056-2981
Yufei Li https://orcid.org/0000-0002-3449-6839
Peiyu Huang https://orcid.org/0000-0003-4226-9369
Minming Zhang https://orcid.org/0000-0003-0145-7558
Zhong Xue https://orcid.org/0000-0003-3207-0890

REFERENCES

Aljabar, P., Heckemann, R. A., Hammers, A., Hajnal, J. V., & Rueckert, D. (2009). Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. *NeuroImage*, *46*(3), 726–738.

Antipov, G., Baccouche, M., & Dugelay, J. L. (2017). Face aging with conditional generative adversarial networks. IEEE International Conference on Image Processing (ICIP), 2089–2093.

Artaechevarria, X., Munoz-Barrutia, A., & Ortiz-de-Solórzano, C. (2009). Combination strategies in multi-atlas image segmentation: Application to brain MR data. *IEEE Transactions on Medical Imaging*, *28*(8), 1266–1277.

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *NeuroImage*, *26*(3), 839–851.

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*(1), 26–41.

Balakrishnan, G., Zhao, A., Sabuncu, M. R., Guttag, J., & Dalca, A. V. (2018). An unsupervised learning model for deformable medical image registration. Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Balakrishnan, G., Zhao, A., Sabuncu, M. R., Guttag, J., & Dalca, A. V. (2019). Voxelmorph: A learning framework for deformable medical image registration. *IEEE Transactions on Medical Imaging*, *38*(8), 1788–1800.

Baloch, S., & Davatzikos, C. (2009). Morphological appearance manifolds in computational anatomy: Groupwise registration and morphological analysis. *NeuroImage*, *45*(1), S73–S85.

Baloch, S., Verma, R., & Davatzikos, C. (2007). An anatomical equivalence class based joint transformation-residual descriptor for morphological analysis. Paper presented at the Biennial International Conference on Information Processing in Medical Imaging.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig, Germany: Verlag von Johann Ambrosius Barth.

Cao, X., Yang, J., Wang, L., Xue, Z., Wang, Q., & Shen, D. (2018). Deep learning based inter-modality image registration supervised by intra-modality similarity. Paper presented at the International Workshop on Machine Learning in Medical Imaging.

Cao, X., Yang, J., Zhang, J., Nie, D., Kim, M., Wang, Q., & Shen, D. (2017). Deformable image registration based on similarity-steered CNN regression. Paper presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention.

Cao, Z., Shi, F., Xu, Q., Liu, G., Sun, T., Xing, X., … Shen, D. (2021). Diagnosis of hippocampal sclerosis from clinical routine head MR images using structure-constrained super-resolution network. Paper presented at the International Workshop on Machine Learning in Medical Imaging.

Chen, H., Dou, Q., Yu, L., Qin, J., & Heng, P.-A. (2018). VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images. *NeuroImage*, *170*, 446–455.

Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-net: Learning dense volumetric segmentation from sparse annotation. Paper presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention.

de Vos, B. D., Berendsen, F. F., Viergever, M. A., Sokooti, H., Staring, M., & Išgum, I. (2019). A deep learning framework for unsupervised affine and deformable image registration. *Medical Image Analysis*, *52*, 128–143.

Dima, D., Modabbernia, A., Papachristou, E., Doucet, G. E., Agartz, I., Aghajani, M., … Alpert, K. I. (2022). Subcortical volumes across the

lifespan: Data from 18,605 healthy individuals aged 3–90 years. *Human Brain Mapping*, *43*(1), 452–469.

Dittrich, E., Raviv, T. R., Kasprian, G., Donner, R., Brugger, P. C., Prayer, D., & Langs, G. (2014). A spatio-temporal latent atlas for semi-supervised learning of fetal brain segmentations and morphological age estimation. *Medical Image Analysis*, *18*(1), 9–21.

Eppenhof, K. A., & Pluim, J. P. (2018). Pulmonary CT registration through supervised learning with convolutional neural networks. *IEEE Transactions on Medical Imaging*, *38*(5), 1097–1105.

Esteban, O., Blair, R. W., Nielson, D. M., Varada, J. C., Marrett, S., Thomas, A. G., … Gorgolewski, K. J. (2019). Crowdsourced MRI quality metrics and expert quality annotations for training of humans and machines. *Scientific Data*, *6*(1), 1–7.

Evans, A. C., Collins, D. L., Mills, S., Brown, E. D., Kelly, R. L., & Peters, T. M. (1993). 3D statistical neuroanatomical models from 305 MRI volumes. Paper presented at the 1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference.

Ferrante, E., Oktay, O., Glocker, B., & Milone, D. H. (2018). On the adaptability of unsupervised CNN-based deformable image registration to unseen image domains. Paper presented at the International Workshop on Machine Learning in Medical Imaging.

Fischl, B. (2012). FreeSurfer. *NeuroImage*, *62*(2), 774–781.

Fisher, R. A. (1992). Statistical methods for research workers. In *Breakthroughs in statistics* (pp. 66–70). Springer.

Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., McKinstry, R. C., Collins, D. L., & Brain Development Cooperative Group. (2011). Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, *54*(1), 313–327.

Frangou, S., Modabbernia, A., Williams, S. C., Papachristou, E., Doucet, G. E., Agartz, I., … Alnæs, D. (2022). Cortical thickness across the lifespan: Data from 17,075 healthy individuals aged 3–90 years. *Human Brain Mapping*, *43*(1), 431–451.

Friston, K. J. (2003). Statistical parametric mapping. *In: Kötter, R. (eds) Neuroscience Databases. Springer, Boston, MA.* 237–250.

Fu, Y., Lei, Y., Wang, T., Curran, W. J., Liu, T., & Yang, X. J. P. (2020). Deep learning in medical image registration: A review. *Physics in Medicine Biology*, *65*(20), 20TR01.

Gu, D., Cao, X., Ma, S., Chen, L., Liu, G., Shen, D., & Xue, Z. (2020). Pairwise and group-wise deformation consistency in deep registration network. Paper presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention.

Habas, P. A., Kim, K., Corbett-Detig, J. M., Rousseau, F., Glenn, O. A., Barkovich, A. J., & Studholme, C. (2010). A spatiotemporal atlas of MR intensity, tissue probability and shape of the fetal brain with application to segmentation. *NeuroImage*, *53*(2), 460–470.

Hamm, J., Davatzikos, C., & Verma, R. (2009). Efficient large deformation registration via geodesics on a learned manifold of images. Paper presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention.

Han, M., Yao, G., Zhang, W., Mu, G., Zhan, Y., Zhou, X., & Gao, Y. (2019). Segmentation of CT thoracic organs by multi-resolution VB-nets. Paper presented at the SegTHOR Challenge@ISBI.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Hedman, A. M., van Haren, N. E., Schnack, H. G., Kahn, R. S., & Hulshoff Pol, H. E. (2012). Human brain changes across the life span: A review of 56 longitudinal magnetic resonance imaging studies. *Human Brain Mapping*, *33*(8), 1987–2002.

Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., & Evans, A. C. (1998). Enhancement of MR images using registration for signal averaging. *Journal of Computer Assisted Tomography*, *22*(2), 324–333.

Hu, X., Kang, M., Huang, W., Scott, M. R., Wiest, R., & Reyes, M. (2019). Dual-stream pyramid registration network. Paper presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention.

Hua, R., Huo, Q., Gao, Y., Sui, H., Zhang, B., Sun, Y., … Shi, F. (2020). Segmenting brain tumor using cascaded V-nets in multimodal MR images. *Frontiers in Computational Neuroscience*, *14*, 9.

Jaderberg, M., Simonyan, K., & Zisserman, A. (2015). Spatial transformer networks. Paper presented at the Advances in Neural Information Processing Systems.

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790.

Jia, H., Wu, G., Wang, Q., & Shen, D. (2010). ABSORB: Atlas building by self-organized registration and bundling. *NeuroImage*, *51*(3), 1057–1070.

Jia, H., Yap, P.-T., Wu, G., Wang, Q., & Shen, D. (2011). Intermediate templates guided groupwise registration of diffusion tensor images. *NeuroImage*, *54*(2), 928–939.

Joshi, S., Davis, B., Jomier, M., & Gerig, G. (2004). Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*, *23*, S151–S160.

Kim, M., Wu, G., Yap, P.-T., & Shen, D. (2011). A general fast registration framework by learning deformation–appearance correlation. *IEEE Transactions on Image Processing*, *21*(4), 1823–1833.

Kingma, D. P., & Welling, M. (2019). An introduction to variational autoencoders. *Foundations and Trends in Machine Learning*, *12*(4), 307–392.

Klein, A., & Hirsch, J. (2005). Mindboggle: A scatterbrained approach to automate brain labeling. *NeuroImage*, *24*(2), 261–280.

Kruskal, J. B. (1956). On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society*, *7*(1), 48–50.

Kuklisova-Murgasova, M., Aljabar, P., Srinivasan, L., Counsell, S. J., Doria, V., Serag, A., … Edwards, A. D. (2011). A dynamic 4D probabilistic atlas of the developing brain. *NeuroImage*, *54*(4), 2750–2763.

Lan, R., Sun, L., Liu, Z., Lu, H., Su, Z., Pang, C., & Luo, X. (2020). Cascading and enhanced residual networks for accurate single-image super-resolution. *IEEE Transactions on Cybernetics*, *51*(1), 115–125.

Learned-Miller, E. G. (2005). Data driven image models through continuous joint alignment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*(2), 236–250.

Liang, P., Shi, L., Chen, N., Luo, Y., Wang, X., Liu, K., … Li, K. (2015). Construction of brain atlases based on a multi-center MRI dataset of 2020 Chinese adults. *Scientific Reports*, *5*(1), 1–7.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., … Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60–88.

Liu, G., Cao, Z., Xu, Q., Zhang, Q., Yang, F., Xie, X., … He, Y. (2021). Recycling diagnostic MRI for empowering brain morphometric research–critical & practical assessment on learning-based image super-resolution. *NeuroImage*, *245*, 118687.

Lv, J., Yang, M., Zhang, J., & Wang, X. (2018). Respiratory motion correction for free-breathing 3D abdominal MRI using CNN-based image registration: A feasibility study. *The British Journal of Radiology*, *91*, 20170788.

Mandal, P. K., Mahajan, R., & Dinov, I. D. (2012). Structural brain atlases: Design, rationale, and applications in normal and pathological cohorts. *Journal of Alzheimer's Disease*, *31*(s3), S169–S188.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., … Pike, B. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *356*(1412), 1293–1322.

Munsell, B. C., Temlyakov, A., & Wang, S. (2009). Fast multiple shape correspondence by pre-organizing shape instances. Paper presented at the 2009 IEEE Conference on Computer Vision and Pattern Recognition.

Nielsen, J. A., Zielinski, B. A., Fletcher, P. T., Alexander, A. L., Lange, N., Bigler, E. D., … Anderson, J. (2013). Multisite functional connectivity MRI classification of autism: ABIDE results. *Frontiers in Human Neuroscience*, *7*, 599.

Park, H., Bland, P. H., Hero, A. O., & Meyer, C. R. (2005). Least biased target selection in probabilistic atlas construction. Paper presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention.

Pereira, S., Pinto, A., Oliveira, J., Mendrik, A. M., Correia, J. H., & Silva, C. A. (2016). Automatic brain tissue segmentation in MR images using random forests and conditional random fields. *Journal of Neuroscience Methods*, *270*, 111–123.

Podgórski, P., Bladowska, J., Sasiadek, M., & Zimny, A. (2021). Novel volumetric and surface-based magnetic resonance indices of the aging brain—Does male and female brain age in the same way? *Frontiers in Neurology*, *12*, 645729.

Rohé, M.-M., Datar, M., Heimann, T., Sermesant, M., & Pennec, X. (2017). SVF-net: Learning deformable image registration using shape matching. Paper presented at the International conference on medical image computing and computer-assisted intervention.

Scahill, R. I., Frost, C., Jenkins, R., Whitwell, J. L., Rossor, M. N., & Fox, N. C. (2003). A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. *Archives of Neurology*, *60*(7), 989–994.

Sedghi, A., Luo, J., Mehrtash, A., Pieper, S., Tempany, C. M., Kapur, T., … Wells III, W. M. (2018). Semi-supervised deep metrics for image registration. *arXiv*. Retrieved from https://arxiv.org/abs/1804.01565

Seghers, D., D'Agostino, E., Maes, F., Vandermeulen, D., & Suetens, P. (2004). Construction of a brain template from MR images using state-of-the-art registration and segmentation techniques. Paper presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention.

Serag, A., Aljabar, P., Ball, G., Counsell, S. J., Boardman, J. P., Rutherford, M. A., … Rueckert, D. (2012). Construction of a consistent high-definition spatio-temporal atlas of the developing brain using adaptive kernel regression. *NeuroImage*, *59*(3), 2255–2265.

Shen, D., & Davatzikos, C. (2002). HAMMER: Hierarchical attribute matching mechanism for elastic registration. *IEEE Transactions on Medical Imaging*, *21*(11), 1421–1439.

Shi, F., Fan, Y., Tang, S., Gilmore, J. H., Lin, W., & Shen, D. (2010). Neonatal brain image segmentation in longitudinal MRI studies. *NeuroImage*, *49*(1), 391–400.

Shi, F., Wang, L., Wu, G., Li, G., Gilmore, J. H., Lin, W., & Shen, D. (2014). Neonatal atlas construction using sparse representation. *Human Brain Mapping*, *35*(9), 4663–4677.

Shi, F., Yap, P.-T., Fan, Y., Gilmore, J. H., Lin, W., & Shen, D. (2010). Construction of multi-region-multi-reference atlases for neonatal brain MRI segmentation. *NeuroImage*, *51*(2), 684–693.

Shi, F., Yap, P.-T., Wu, G., Jia, H., Gilmore, J. H., Lin, W., & Shen, D. (2011). Infant brain atlases from neonates to 1-and 2-year-olds. *PLoS One*, *6*(4), e18746.

Sokooti, H., De Vos, B., Berendsen, F., Lelieveldt, B. P., Išgum, I., & Staring, M. (2017). Nonrigid image registration using multi-scale 3D convolutional neural networks. Paper presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention.

Stergios, C., Mihir, S., Maria, V., Guillaume, C., Marie-Pierre, R., Stavroula, M., & Nikos, P. (2018). Linear and deformable image registration with 3d convolutional neural networks. In *Image analysis for moving organ, breast, and thoracic images* (pp. 13–22). Springer.

Stollenga, M. F., Byeon, W., Liwicki, M., & Schmidhuber, J. (2015). Parallel multi-dimensional lstm, with application to fast biomedical volumetric image segmentation. *arXiv preprint arXiv:07452*.

Talairach, J. (1988). *Co-planar stereotaxic atlas of the human brain-3-dimensional proportional system: An approach to cerebral imaging*. Thieme.

Tang, Y., Hojatkashani, C., Dinov, I. D., Sun, B., Fan, L., Lin, X., … Toga, A. W. (2010). The construction of a Chinese MRI brain atlas: A morphometric comparison study between Chinese and Caucasian cohorts. *NeuroImage*, *51*(1), 33–41.

Uzunova, H., Wilms, M., Handels, H., & Ehrhardt, J. (2017). Training CNNs for image registration from few samples with model-based data augmentation. Paper presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention.

van Opbroek, A., van der Lijn, F., & de Bruijne, M. (2013). Automated brain-tissue segmentation by multi-feature SVM classification. Paper presented at the MICCAI Grand Challenge on MR Brain Image Segmentation (MRBrainS13).

Vercauteren, T., Pennec, X., Perchant, A., & Ayache, N. (2009). Diffeomorphic demons: Efficient non-parametric image registration. *NeuroImage*, *45*(1), S61–S72.

Wang, L., Gao, Y., Shi, F., Li, G., Gilmore, J. H., Lin, W., & Shen, D. (2015). LINKS: Learning-based multi-source integration framework for segmentation of infant brain images. *NeuroImage*, *108*, 160–172.

Wang, Q., Wu, G., Yap, P.-T., & Shen, D. (2010). Attribute vector guided groupwise registration. *NeuroImage*, *50*(4), 1485–1496.

Wei, J., Shi, F., Cui, Z., Pan, Y., Xia, Y., & Shen, D. (2021). Consistent segmentation of longitudinal brain MR images with spatio-temporal constrained networks. Paper presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention, Strasbourg, France.

Wu, G., Jia, H., Wang, Q., & Shen, D. (2011). SharpMean: Groupwise registration guided by sharp mean image and tree-based registration. *NeuroImage*, *56*(4), 1968–1981.

Wu, G., Wang, Q., Jia, H., & Shen, D. (2012). Feature-based groupwise registration by hierarchical anatomical correspondence detection. *Human Brain Mapping*, *33*(2), 253–271.

Xiao, B., Cheng, X., Li, Q., Wang, Q., Zhang, L., Wei, D., … Lu, G. (2019). Weakly supervised confidence learning for brain MR image dense parcellation. Paper presented at the International Workshop on Machine Learning in Medical Imaging.

Yang, G., Zhou, S., Bozek, J., Dong, H.-M., Han, M., Zuo, X.-N., … Gao, J.-H. (2020). Sample sizes and population differences in brain template construction. *NeuroImage*, *206*, 116318.

Yang, X. (2017). *Uncertainty quantification, image synthesis and deformation prediction for image registration*. (PhD thesis). The University of North Carolina.

Zacharaki, E. I., Dinggang Shen, Seung-Koo Lee, & Davatzikos, C. (2008). ORBIT: A multiresolution framework for deformable registration of brain tumor images. *IEEE Transactions on Medical Imaging*, *27*(8), 1003–1017.

Zhang, Y., Shi, F., Wu, G., Wang, L., Yap, P.-T., & Shen, D. (2016). Consistent spatial-temporal longitudinal atlas construction for developing infant brains. *IEEE Transactions on Medical Imaging*, *35*(12), 2568–2577.

Zitova, B., & Flusser, J. (2003). Image registration methods: A survey. *Image and Vision Computing*, *21*(11), 977–1000.

Zuo, X.-N., Anderson, J. S., Bellec, P., Birn, R. M., Biswal, B. B., Blautzik, J., … Castellanos, F. X. (2014). An open science resource for establishing reliability and reproducibility in functional connectomics. *Scientific Data*, *1*(1), 1–13.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Gu, D., Shi, F., Hua, R., Wei, Y., Li, Y., Zhu, J., Zhang, W., Zhang, H., Yang, Q., Huang, P., Jiang, Y., Bo, B., Li, Y., Zhang, Y., Zhang, M., Wu, J., Shi, H., Liu, S., He, Q., Zhang, Q., Zhang, X., Wei, H., Liu, G., Xue, Z., Shen, D., & the Consortium of Chinese Brain Molecular and Functional Mapping (CBMFM) (2023). An artificial-intelligence-based age-specific template construction framework for brain structural analysis using magnetic resonance images. *Human Brain Mapping*, *44*(3), 861–875. https://doi.org/10.1002/hbm.26126