

APPLIED SCIENCES AND ENGINEERING

SynthSR: A public AI tool to turn heterogeneous clinical brain scans into high-resolution T1-weighted images for 3D morphometry

Juan E. Iglesias^{1,2,3*}, Benjamin Billot², Yaël Balbastre¹, Colin Magdamo⁴, Steven E. Arnold⁴, Sudeshna Das⁴, Brian L. Edlow^{1,4,5}, Daniel C. Alexander², Polina Golland³, Bruce Fischl^{1,3}

Every year, millions of brain magnetic resonance imaging (MRI) scans are acquired in hospitals across the world. These have the potential to revolutionize our understanding of many neurological diseases, but their morphometric analysis has not yet been possible due to their anisotropic resolution. We present an artificial intelligence technique, “SynthSR,” that takes clinical brain MRI scans with any MR contrast (T1, T2, etc.), orientation (axial/coronal/sagittal), and resolution and turns them into high-resolution T1 scans that are usable by virtually all existing human neuroimaging tools. We present results on segmentation, registration, and atlas of >10,000 scans of controls and patients with brain tumors, strokes, and Alzheimer’s disease. SynthSR yields morphometric results that are very highly correlated with what one would have obtained with high-resolution T1 scans. SynthSR allows sample sizes that have the potential to overcome the power limitations of prospective research studies and shed new light on the healthy and diseased human brain.

INTRODUCTION

Neuroimaging with MRI is one of the most useful tools available to study the human brain in vivo. Open-source neuroimaging software packages like FreeSurfer (1), FSL (2), SPM (3), and AFNI (4) have enabled researchers around the world to conduct brain studies in an automated fashion, e.g., to characterize brain structure and function in healthy aging and in diseases like Alzheimer’s. These tools have also increased reproducibility of results, particularly when combined with publicly available datasets such as Alzheimer’s Disease Neuroimaging Initiative (ADNI) (5), the Human Connectome Project (6), or the UK Biobank (7).

The automated processing methods in current neuroimaging tools require magnetic resonance imaging (MRI) scans acquired with high, isotropic resolution (typically 1 mm) to minimize errors in three-dimensional (3D) analyses such as segmentation (8–10) or registration (11–16). In addition, several methods have requirements in terms of MR contrast. For example, FreeSurfer requires T1-weighted scans. For this reason, most modern research neuroimaging studies include a structural MRI acquisition that fulfills such resolution and MR contrast requirements—often a 1-mm isotropic, T1-weighted scan acquired with the ubiquitous 3D magnetization prepared - rapid gradient echo (MPRAGE) pulse sequence (17).

However, the vast majority of MRI scans in the world are acquired for clinical purposes and do not satisfy the aforementioned criteria. In the clinic, physicians typically prefer 2D acquisitions that yield fewer slices (e.g., 20 to 30) with large spacing (5 to 7 mm) and

high in-plane resolution (under 1 mm), often acquired with the widespread turbo spin echo (TSE) sequence (18). This type of acquisition reduces the time that is required to inspect the images and is less sensitive to motion artifacts, which is crucial for patients whose neurological diseases make it challenging for them to lie still in the MRI scanner.

The inability to compute morphometric measurements from clinical scans for use in neuroimaging research is a major barrier to progress in this field, as it precludes the analysis of millions of scans that are currently sitting in picture archiving and communication systems (PACS) in hospitals worldwide. For example, approximately 10 million brain MR exams were performed in the United States in 2019 alone (19). These figures are huge compared with even the largest neuroimaging MRI studies and meta-studies (20), and could yield sample sizes with the potential to revolutionize our understanding of neurological conditions and genetic linkages, compared with typically underpowered prospective research studies (21).

Artificial intelligence (AI) techniques—particularly emerging deep machine learning (ML) models for image synthesis and super-resolution (SR)—have the potential to bridge the gap between clinical and research-grade brain MRI scans. Given a scan of a certain (non-T1) MR contrast and low (non-isotropic) resolution, these techniques can be used to generate a new imaging volume that has T1-like contrast (via synthesis) and high, isotropic resolution (via SR).

There are many existing techniques for synthesis and SR of MRI, based both on classical and deep ML methods. Modern SR methods are almost exclusively based on deep learning (22, 23), specifically convolutional neural networks (CNNs). These typically capitalize on large amounts of paired low- and high-resolution images (LR/HR) to learn mappings from the former to the latter (24–26). Further refinement of the output can be achieved by enhancing the architecture with adversarial losses (27), i.e., by trying to fool a discriminator that is trained to separate real HR images from

Copyright © 2023 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).

¹Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. ²Centre for Medical Image Computing, Department of Computer Science, University College London, London, UK. ³Computer Science and Artificial Intelligence Laboratory (CSAIL), Massachusetts Institute of Technology, Cambridge, MA, USA. ⁴Department of Neurology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. ⁵Center for Neurotechnology and Neurorecovery, Massachusetts General Hospital, Boston, MA, USA.

*Corresponding author. Email: jiglesiasgonzalez@mgh.harvard.edu

enhanced LR images (28–30). Paired data are easy to obtain, by taking HR images and downsampling them to obtain LR counterparts.

Similar to SR, modern synthesis methods are based on CNNs seeking to learn a mapping between a source and a target modality—often enhanced with adversaries, too, trained to discriminate real and synthetic images. While obtaining perfectly paired data is more difficult than in SR, this limitation has been mitigated with unpaired techniques based on generative adversarial networks (GANs), which seek to generate synthetic images of the target modality that are difficult to discriminate from real ones by an auxiliary CNN (27). A representative example is the CycleGAN framework (31), which combines GANs with cycle consistency, i.e., mapping an image to a target domain and then back to the source should be close to the identity operator. However, GANs are typically used in combination with supervised losses based on voxel-wise errors, since they underperform when used in isolation (32). Driven by applications like estimation of missing modalities and synthesis of computed tomography (CT) from MRI for attenuation correction in positron emission tomography (PET), many methods have been developed for image synthesis in brain MRI (33–39).

Unfortunately, three main roadblocks have precluded the application of deep learning SR and synthesis techniques to clinical brain MRI data. First, the domain shift: The performance of CNNs quickly decreases when the resolution or MR contrast of the input diverges from the data that the network was trained on (40). This gap is particularly problematic in clinical brain MRI, due to the huge diversity in orientation, resolution, and MR contrast of acquisitions, both within and across centers. Data augmentation (41) and domain adaptation techniques (42) mitigate the problem but have not been able to close the gap. Self-supervised SR techniques exploit the high-frequency content within slices to learn to super-resolve the LR in orthogonal views across slices (43), but cannot tackle the synthesis problem, since they only have access to the target modality. The second obstacle is the need to retrain. Even if training data were available for every possible combination of orientation, contrast, and resolution, different MRI exams with different sets of pulse sequences would require additional runs of training or, at least, domain adaptation. Self-supervised methods also suffer from this limitation. Therefore, there are currently no methods that can be used “out of the box” to process heterogeneous datasets. The third barrier is the modeling of pathology: To the best of our knowledge, no existing SR or synthesis method is robust to the wide variation of pathology that one encounters in a PACS.

Here, we present our neural network SynthSR, which we distribute with FreeSurfer. SynthSR turns a clinical brain MRI scan of any orientation, resolution and contrast into a 1-mm isotropic 3D MPRAGE. This synthetic MPRAGE can be subsequently analyzed with any existing tool for 3D image analysis of brain MRI, e.g., registration or segmentation. SynthSR is an evolution of our previous tool (44), which could be trained to process images of predefined resolution and contrast using synthetic data, yet suffered from the three limitations described above. Our new tool, on the other hand, (i) combines a domain randomization (DR) approach (45) with a generative model of brain MRI to support scans of any resolution and contrast out of the box without retraining, (ii) produces more realistic images via an auxiliary segmentation task, and (iii) handles abnormalities by inpainting them with normal-looking tissue. Because most standard neuroimaging tools like FreeSurfer or FSL

cannot cope with pathology (particularly large lesions like some tumors or strokes), inpainting them with healthy tissue enables direct subsequent analysis with these tools. This approach is common, e.g., in the multiple sclerosis literature, where white matter lesions are filled with white matter–like intensities before 3D morphometry with packages like FreeSurfer or SPM (46).

SynthSR is publicly available and can be easily used by downloading FreeSurfer (<https://surfer.nmr.mgh.harvard.edu/fswiki/DownloadAndInstall>) and executing the command

```
mri_synthsr --i [clinical_scan] --o [synthetic_1mm_isotropic_t1]
```

which only takes a few seconds to run on a graphics processing unit (GPU) or approximately 15 s on a modern desktop computer without a GPU.

RESULTS

Image segmentation and volumetry with existing tools

One of the main use cases of SynthSR is automated volumetry of regions of interest (ROIs) from MRI scans that do not satisfy the requirements of the packages that are often used for this purpose. We used SynthSR to process the 9146 brain MRI scans in the Massachusetts General Hospital (MGH) dataset, which includes nearly uncurated data from 1110 patients with neurology visits at MGH; we note that we left out scans with more than three dimensions (e.g., diffusion-weighted MRI) or with intracranial volume (ICV) under 1.1 liters (typically with partial field of view); further details can be found in Materials and Methods.

Examples of the synthetic outputs are shown in Fig. 1. We segmented these synthetic images with FreeSurfer and compared the

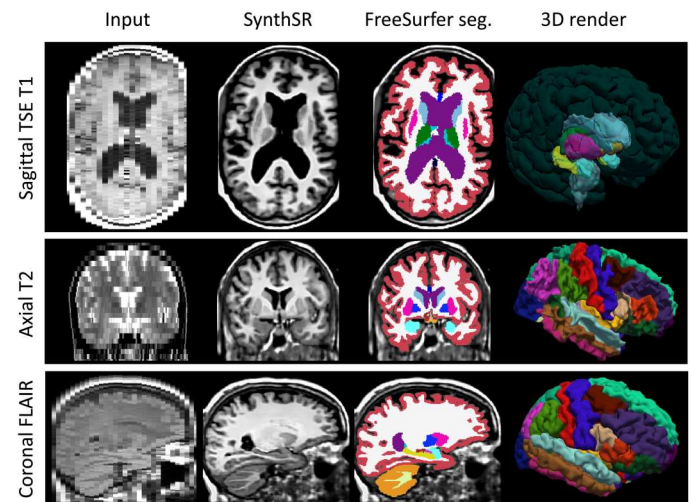


Fig. 1. Examples of inputs and outputs of SynthSR from the MGH dataset. Each of the inputs has been acquired with a different orientation (axial, sagittal, and coronal), slice spacing (6, 5, and 4 mm), and MR contrast (turbo spin echo T1, T2, FLAIR); we visualize them in orthogonal view to illustrate their low-resolution out of plane. SynthSR produces a synthetic 1-mm isotropic MPRAGE in all cases, which is compatible with nearly every existing brain MR image analysis method. The third column shows the automated segmentation of the synthetic MPRAGE obtained with FreeSurfer, and the last column shows three-dimensional (3D) renderings of (top to bottom) the subcortical segmentation, parcellated white matter surface, and parcellated pial surface. We emphasize that all images were processed with the same neural network using SynthSR out of the box, without retraining.

volumes derived from these segmentations with ground-truth estimates for a subset of 435 scans from 41 patients who also had a 1-mm isotropic T1-weighted scan. Correlations at the scan and subject level (computed as the median of the volumes across available scans within a session) are shown in Table 1. The correlations at the single scan level are strong (~0.8) for tissue classes (white matter, cortical gray matter, and subcortical gray matter), nearly perfect for the ventricles, and moderate to strong for individual subcortical ROIs—ranging from 0.47 for the pallidum (which is very difficult to segment in T1 due to the lack of contrast) to 0.76 for the hippocampus. When the measurements from a single session are aggregated into a single estimate, these correlations increase considerably, becoming very strong (~0.9) for nearly every tissue class and subcortical ROI, except for the putamen and pallidum—for the aforementioned reasons.

To study the contribution of the auxiliary segmentation task to the performance SynthSR, we performed an ablation study where the auxiliary segmentation task was disregarded during training. When this alternative model was used, correlations decreased for almost all ROIs, especially those with faint or convoluted boundaries (e.g., cortex, amygdala, and pallidum). This result highlights the importance of the segmentation loss for accurately synthesizing such brain regions.

To study the performance of SynthSR as a function of resolution, we stratified the correlations by resolution using groups of scans with similar slice spacing. The results are shown in Table 2 and show a clear pattern of decrease in correlation with growing spacing. Some larger ROIs like the ventricles or cortical gray matter are more robust to slice spacing than smaller ROIs like the amygdala or pallidum. We also note that aggregating results at the subject level (i.e., bottom row of Table 1) yields, on average, higher correlations than the scans with smaller spacing on their own (top row of Table 2).

We then used the segmentations from the MGH dataset to test whether we could replicate well-established atrophy patterns due to aging. Figure 2 shows individual points (at the scan level) and regressed median trajectories with confidence intervals for different tissue classes and subcortical ROIs. Trajectories at the subject level are shown in fig. S1. The trajectories in Fig. 2 are remarkably similar to those from a recent meta-analysis with more than 100,000 HR isotropic scans (47): They correctly capture the peak of the white matter at about 30 years, the earlier decline of the gray matter, and the highly nonlinear trajectory of the ventricles. Despite using clinical scans with large spacing, our method also produces trajectories

for subcortical ROIs that are highly consistent with previously published studies relying on thousands of 1-mm isotropic T1 scans (48, 49), showing, e.g., earlier atrophy of the thalamus compared with hippocampus or amygdala.

Detecting disease-induced atrophy: Alzheimer's disease volumetry

The following experiment seeks to illustrate the ability of SynthSR to preserve the effects of disease on the shapes and volumes of ROIs when processing scans with large slice spacing while demonstrating its compatibility with packages other than FreeSurfer. For this purpose, we combined SynthSR and FSL to detect hippocampal atrophy in Alzheimer's disease (AD). Using this application has two advantages. First, hippocampal atrophy is a well-established biomarker of this type of dementia (50–52), so we know what results to expect. Second, we can use publicly available scans from the ADNI dataset (5), which has the advantage of including 5-mm axial fluid attenuated inversion recovery (FLAIR) scans (which we feed to SynthSR) and 1-mm isotropic MPRAGE scans (which we use as ground truth).

In this experiment, we used a sample with 50 randomly selected AD cases and 50 randomly selected controls (47 males, 53 females; aged 72.9 ± 7.6 years). This sample size is representative of modest-sized neuroimaging research studies and ensures that the analysis is not overpowered to the point that even poor hippocampal segmentations lead to very small *P* values. We used FSL to compute segmentations and derive volumes from the synthetic scans estimated from the FLAIR acquisitions; the volumes estimated from the real MPRAGEs directly with FSL were used as ground truth. We emphasize that hippocampal volumetry with this dataset is a very challenging task: Because the major axis of the hippocampus is approximately parallel to the axial plane, this ROI is visible in very few slices (often just two to four).

Figure 3 (A to D) shows the output of SynthSR and the subsequent FSL segmentation for a sample scan. Qualitatively, SynthSR recovers many of the missing high-frequency details and enables accurate segmentation with FSL. Figure 3E shows a scatterplot for the ground truth and estimated hippocampal volumes (note that we excluded four outliers for which the FSL segmentation failed on the real 1-mm T1 scan). The plot reveals a strong correlation ($\rho = 0.83$, $P < 10^{-24}$) between the volumes estimated from the real T1s and the synthetic MPRAGEs. The latter have a slight positive bias due to the smoothing that is introduced by SynthSR when interpolating between distant MRI slices.

Table 1. Correlation with ground-truth volumes. Pearson correlations between volumes obtained with FreeSurfer from 1-mm isotropic T1s and from the synthetic MPRAGEs computed with SynthSR from clinical acquisitions of heterogeneous orientation, resolution, and contrast. The correlations at the scan level are computed with and without the auxiliary segmentation loss to analyze the impact of this component of our method. The correlations at the subject level use volume estimates computed as the median across available scans for each subject (using the full SynthSR model, i.e., with the segmentation loss). All correlations are strongly significant ($P < 10^{-7}$ when using the full model).

Region level	White matter	Cortical gray matter	Subcortical gray matter	Ventricles	Hippocampus	Amygdala	Thalamus	Putamen	Pallidum
Scan level ($n = 435$)	0.79	0.83	0.77	0.99	0.76	0.60	0.72	0.60	0.47
Scan level (ablated segmentation task)	0.79	0.79	0.76	0.99	0.74	0.54	0.69	0.56	0.40
Subject level ($n = 41$)	0.91	0.93	0.91	0.99	0.89	0.90	0.90	0.75	0.72

Table 2. Performance as a function of slice spacing. Pearson correlations between volumes obtained with FreeSurfer from 1-mm isotropic T1s and from the synthetic MPRAGEs were computed with SynthSR from clinical acquisitions of heterogeneous orientation and contrast as a function of slice spacing. All correlations are strongly significant ($P < 10^{-3}$).

Region slice spacing	White matter	Cortical gray matter	Subcortical gray matter	Ventricles	Hippocampus	Amygdala	Thalamus	Putamen	Pallidum
Up to 4 mm ($n = 197$)	0.92	0.84	0.90	0.99	0.89	0.80	0.92	0.73	0.67
Between 4 and 6 mm ($n = 110$)	0.71	0.84	0.71	0.99	0.69	0.55	0.63	0.50	0.41
More than 6 mm ($n = 128$)	0.68	0.80	0.64	0.99	0.56	0.37	0.59	0.49	0.27

Because the hippocampal volumes of both AD subjects and controls (as estimated from the real T1s) did not follow Gaussian distributions— $P < 0.05$ for a Shapiro-Wilk test (53)—we used nonparametric statistics to compare AD versus controls. In this setting, we compared SynthSR against (i) analysis of the real 1-mm MPRAGE scans, which provides a ceiling for the performance of the synthesis methods and enables quantification of the gap with respect to such ceiling; (ii) a fully supervised U-net, trained with spatially aligned pairs of real scans, which provides a ceiling for the performance of synthesis-based methods; (iii) an ablation of the lesion modeling component of SynthSR, which enables us to assess whether this building block has an impact on the ability of the tool to detect disease-induced atrophy; and (iv) our previous version of the tool (44)—plus the segmentation loss (for fair comparison)—which enables us to assess whether DR in SynthSR incurs a decrease in performance with respect to training with simulations of the specific resolution and contrast of the FLAIR scans.

Figure 3F shows the median of the two groups, the P value for a nonparametric Wilcoxon rank sum test (54), and the area under the receiver operating characteristic curve (AUROC), which quantifies the separation between the two classes in a nonparametric setting. Despite the sparse slices in an orientation almost parallel to the major axis of the hippocampus (as mentioned above, it often appears in only two to four slices), SynthSR yields very strong discriminative power between the two groups, almost as much as the real 1-mm isotropic T1s. Our approach only loses 12% of the separation of the medians (481 versus 543 mm^3) and five AUROC points (0.76 versus 0.81) and provides a very strong P value ($\sim 10^{-6}$ versus $\sim 10^{-8}$). Compared with the fully supervised U-net, which has access to the real FLAIR intensities in training, SynthSR only loses two AUROC points. Ablating the lesion modeling component of SynthSR slightly worsens the results (AUROC = 0.75); this result supports the hypothesis that the inpainting step (which is crucial in many applications of SynthSR, as shown in the rest of experiments below) does not have a negative impact on the sensitivity of the method to disease-induced atrophy. Last, training on synthetic images that are simulated to resemble the contrast and resolution of the 5-mm axial FLAIR scans works slightly worse than fully randomizing resolution and contrast during training. We hypothesize that this is because DR mitigates the inaccuracies of the model in terms of, e.g., slice selection profiles, noise, or bias field.

Improving registration of clinical scans: Application to MRI of brain tumors

SynthSR can also improve the accuracy of image registration of clinical brain MRI scans. Image registration (55), i.e., spatial alignment of pairs of images, has found wide application in neuroimaging, e.g., in areas like longitudinal analysis (56), fusion of multimodality scans (57), or creation of population atlases (58). One important application of registration is the spatial mapping of neuroanatomical correspondences in preoperative and follow-up scans of patients with glioma, as finding features in the former that can predict tumor infiltration and recurrence is crucial for guiding treatment (59). However, this registration can be difficult due to differences in orientation, resolution, MR contrast, and tumor size and appearance of the preoperative and follow-up scans—a problem that SynthSR can mitigate by synthesizing 1-mm isotropic MPRAGE images with inpainted tumors.

We applied SynthSR to Brain Tumor Sequence Registration (BraTS-Reg), a recently released dataset of multimodal brain scans of patients with glioma, acquired pre- and postoperatively, which includes manual annotations of corresponding landmarks (60). The dataset includes scans from 140 patients and 1252 landmarks. We note that we detected some outliers in the landmarks and are working with the authors of the article to amend them. The results presented here are for a subset of 1075 landmarks that passed our quality control. We provide the list of landmarks that passed quality control in the Supplementary Materials.

The set of modalities is fixed for all subjects and time points: T1, T1 with contrast enhancement (T1CE), T2, and FLAIR; however, the resolution varies across contrasts and time points, as BraTS-Reg is a “multi-institutional dataset consisting of scans acquired under routine clinical conditions, and hence reflecting very heterogeneous acquisition equipment and protocols, affecting the image properties” (60). An example of the application of SynthSR to this dataset is shown in Fig. 4 (A and B), where our method successfully upscales the axial FLAIR scans to 1-mm isotropic resolution while synthesizing MPRAGE contrast and inpainting abnormalities.

We then used the well-established neuroimaging registration package NiftyReg to register the preoperative and follow-up scans. NiftyReg implements a robust affine alignment method based on block matching (61) as well as a nonlinear diffeomorphic registration model based on stationary velocity fields parameterized by grids of control points (62). We registered the scans with the local normalized cross-correlation metric (which is commonly used in MRI due to its robustness to bias fields) and three different

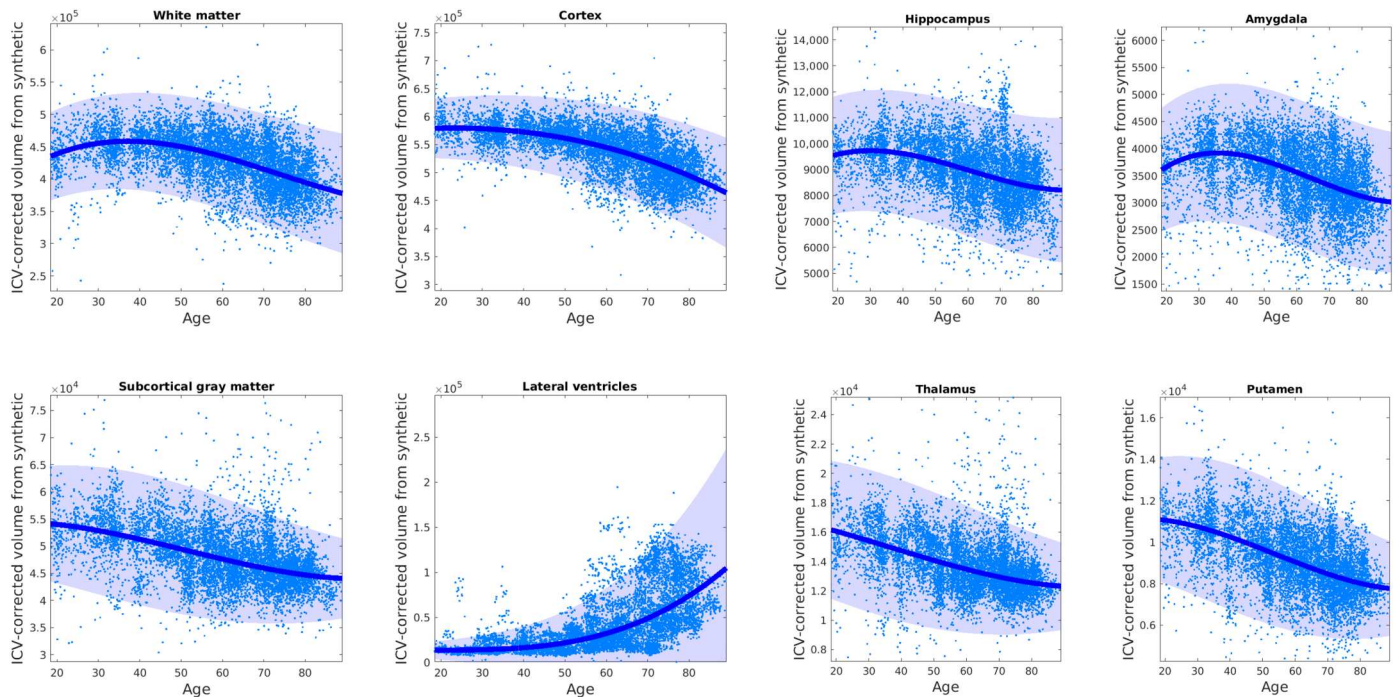


Fig. 2. Brain volume aging trajectories derived from clinical data using SynthSR combined with FreeSurfer. Clinical scans were processed with SynthSR and then segmented with FreeSurfer to obtain ROI volumes as well as an estimate of the intracranial volume (ICV). The ROI volumes were corrected by ICV and sex and regressed against age using a Laplace distribution with location and scale modeled with a spline with six knots. The Laplace distribution provides a more robust fit than a Gaussian. The median trajectory is overlaid on the individual (ICV- and gender-corrected) volumes; the 95% confidence interval is shaded. The volumes are computed at the scan level; trajectories computed at the subject level are shown in fig. S1.

control point spacings (5, 10, and 20 mm), which model different levels of freedom of the nonlinear transform, i.e., different compromises between accuracy (lower spacing) and robustness (higher spacing). Default values were used for all other parameters.

When the original scans are registered directly with NiftyReg, the differences in resolution and appearance of the tumor area (Fig. 4, A and B) make registration difficult. NiftyReg erroneously tries to match tumor boundaries and other partial voluming artifacts, thus yielding highly convoluted deformation fields (Fig. 4, C and D). SynthSR mitigates this problem by computing the registration from synthetic images of isotropic resolution and similar appearance (Fig. 4, E and F), which yields much more regular deformation fields (Fig. 4, G and H). This approach reduces the

average landmark error after registration for every combination of control point spacings and MR contrasts, between 10 and 20% (Fig. 4C). We note that the SD of the error increased for the T1 scans but decreased for the T1CE, T2, and FLAIR scans. In some cases, SynthSR’s robustness enables us to obtain better results with more aggressive (lower) control point spacings, compared with using the original images (e.g., for T1CE or FLAIR images).

We also applied SynthSR to the original BraTS dataset (63), which consists of 1251 cases with manually traced low- and high-grade gliomas, and the same four MR contrasts as BraTS-Reg. Using the synthetic MPRAGEs to compute registrations, we estimated an unbiased atlas (64) of patients with glioma with the four MR contrasts, as well as probabilistic maps of the manually labeled tumor

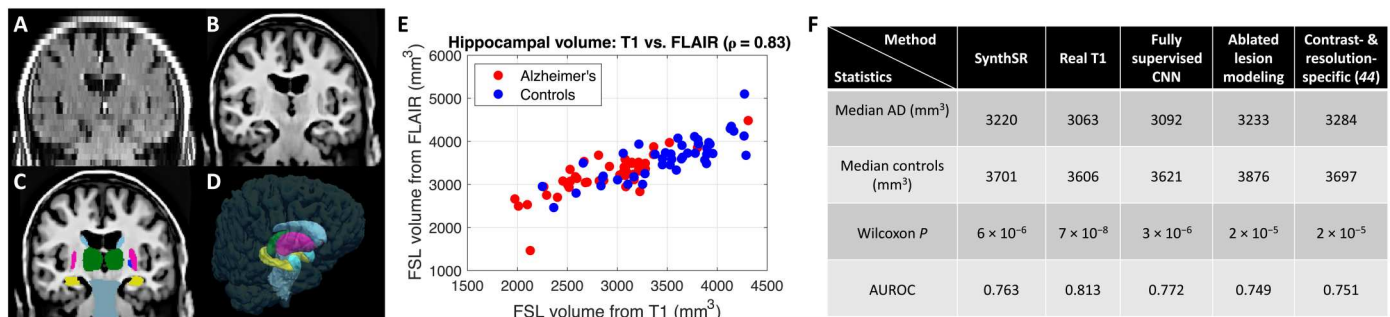


Fig. 3. Hippocampal volumetry of AD with SynthSR and FSL. (A) Coronal view of sample 5-mm axial FLAIR scan. (B) SynthSR. (C) Segmentation of (B) with FSL. (D) 3D rendering of (C). (E) Scatterplot of hippocampal volumes computed with FSL from the 1-mm isotropic T1 scans and from the synthetic MPRAGEs (Pearson correlation: 0.83, $P < 10^{-24}$). (F) Nonparametric statistics for hippocampal volume of AD versus controls: medians, P value for Wilcoxon rank sum test, and AUROC.

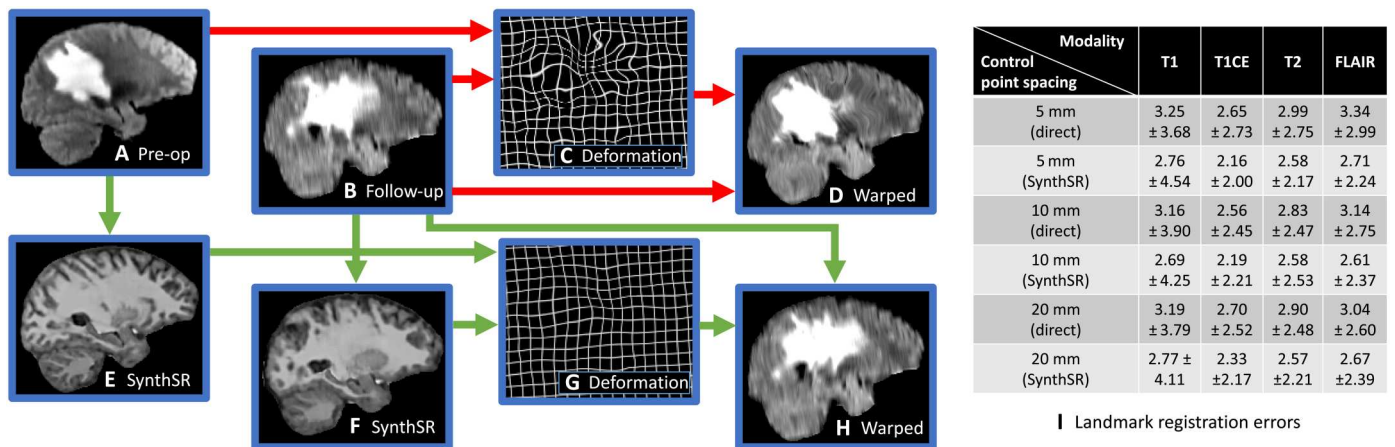


Fig. 4. Alignment of preoperative and follow-up brain MRI scans of patients with glioma (BraTS-Reg dataset) using direct registration (red arrows) and SynthSR (green arrows). (A) Sagittal slice of the FLAIR scan of a sample subject (axial acquisition). (B) Approximately corresponding slice of follow-up scan. (C) Deformation field estimated from (A) and (B) with NiftyReg (5-mm control point spacing). (D) Warped follow-up scan. (E and F) SynthSR output of the preoperative and follow-up scans. (G) Deformation field estimated from the synthetic scans. (H) Warped follow-up scan using the deformation field computed with the synthetic images. (I) Mean and SD of landmark error (in millimeters) for different control point spacings and MR contrasts, registering the scans directly versus registering the output of SynthSR. The *P* values for nonparametric Wilcoxon tests comparing the medians are strongly significant in all cases ($10^{-9} < P < 10^{-3}$).

regions: enhancing tumor, peritumoral edema, and the necrotic and non-enhancing tumor core. Such templates and probability maps are useful in neuroimaging studies for spatial normalization and voxel-wise analysis. The median atlas is shown in Fig. 5. SynthSR overcomes the differences in resolution and the presence of tumors, therefore enabling precise co-registration leading to sharp atlases—particularly in the axial plane, which is the predominant direction of acquisition in the dataset. The accurate alignment also reveals a frontotemporal pattern in the spatial distribution of the gliomas that has been reported in the literature (65) and also captures the tendency of gliomas to infiltrate the white matter.

Segmentation and registration of stroke MRI

Morphometry of brain MRI of patients with stroke is important in the discovery of imaging biomarkers of outcome (66) and to study post-stroke dementia, which is becoming increasingly common as

stroke mortality rates decrease (67). Unfortunately, there are (to the best of our knowledge) no morphometric tools that can readily cope with the abnormal distributions of shape and image intensity distribution caused by stroke, which forces studies to discard many cases that do not pass quality control (68)—not only decreasing sample sizes but also introducing a potential selection bias.

As in the previous experiment, SynthSR can remove differences in resolution, MR contrast, and lesion size and appearance by synthesizing 1-mm MPRAGE scans with inpainted lesions. To evaluate SynthSR in this context, we applied it to ATLAS (69, 70), a publicly available dataset with 655 T1 brain MRI scans of patients with stroke, which includes manually segmented lesion masks. By inpainting stroke lesions, SynthSR enables accurate segmentation with existing packages like FreeSurfer (Fig. 6A). Comparing the ipsi- and contralateral volumes of five representative ROIs (hippocampus, amygdala, thalamus, putamen, and caudate; see Table 3)

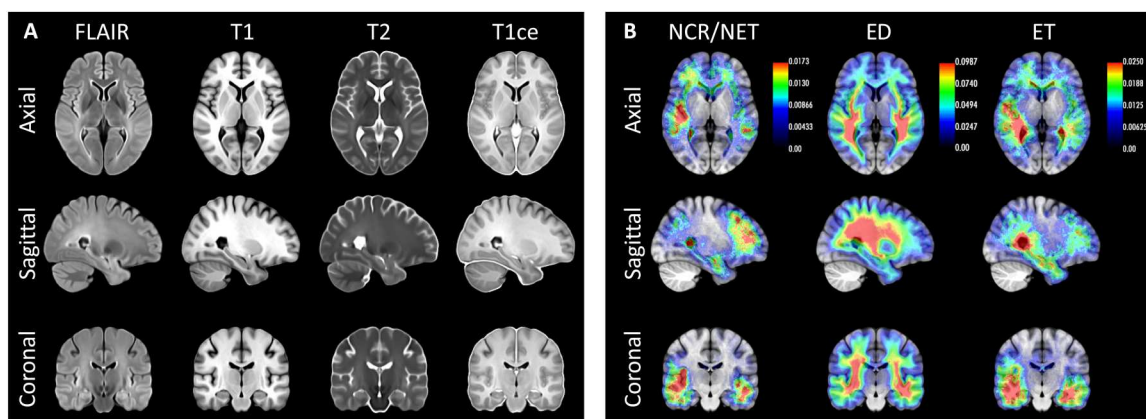


Fig. 5. Population templates and probabilistic maps of tumor subregions for the BraTS dataset. (A) Axial, sagittal, and coronal slices of the group-wise median templates of the four available modalities. (B) Average spatial distribution maps of necrotic and non-enhancing tumor core (NCR/NET), peritumoral edema (ED), and gadolinium-enhancing tumor (ET). The orientation of the images follows radiological convention. The overlaid heatmaps show the probability of observing each tumor class at every voxel, as encoded in the color bars.

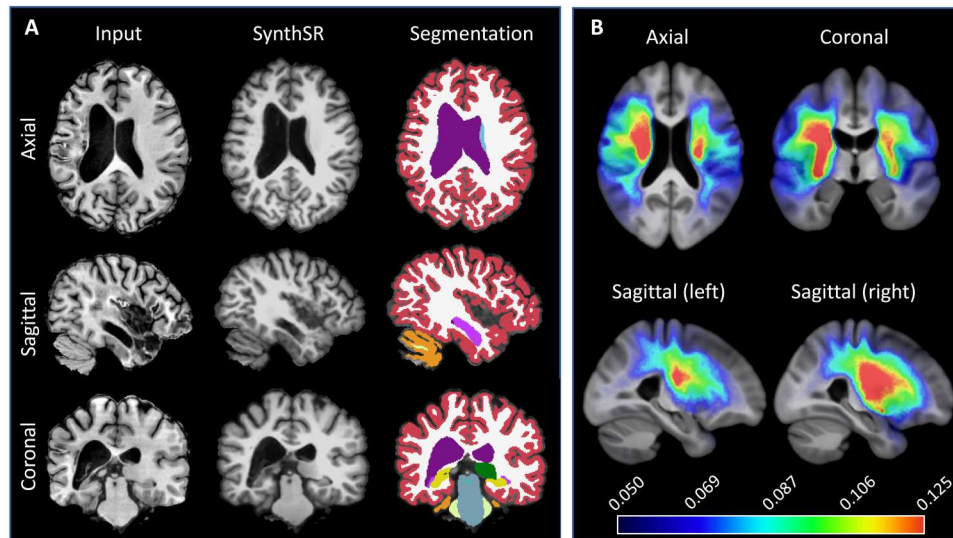


Fig. 6. Application of SynthSR to stroke dataset (ATLAS). (A) Axial, sagittal, and coronal slices of a sample case, along with the output of SynthSR and its segmentation with FreeSurfer. (B) Orthogonal slices of the group-wise median template, along with the average spatial distribution maps of stroke lesions. As in previous figures, the orientation of the images follows radiological convention.

reveals asymmetry patterns that are consistent with the literature, e.g., no asymmetry in the hippocampus (71), but strong asymmetry in the thalamus (72) and the other ROIs. Table 3 also shows the correlation between the volume of the stroke lesion and that of the ROIs in the ipsi- and contralateral sides. The hippocampus shows no statistically significant correlation on either side, while the other ROIs display weak negative correlation on the ipsilateral side but no significant correlation on the contralateral side. The strength of these correlations ($r \sim -0.4$) is consistent with values reported in other studies, e.g., $r = -0.3$ in (73).

SynthSR can also improve registration of these images, which finds application in, e.g., construction of templates for spatial normalization and voxel-wise analysis, for longitudinal analysis, or to correct for edema (74). We performed the atlas-building experiment from the previous section and obtained the templates in Fig. 6B. The average lesion is centered in the basal ganglia and deep white matter. One possible explanation would be an overrepresentation of small vessel strokes in the ATLAS dataset, since these are prevalent in the thalamo-capsular region. The lesion map also shows that strokes are on average larger on the right hemisphere, since the heatmap is stronger and the total number of left and right hemispheric strokes is approximately the same. Once again, further assessment will be required to determine whether this is an actual effect or the consequence of a selection bias, e.g., large lesions on the left hemisphere may affect speech and lead to reduced enrollment rates in ATLAS.

DISCUSSION

We present a publicly available AI technique (SynthSR) that can transform any clinical brain MRI scan into a synthetic 1-mm MPRAGE that is compatible with practically every image analysis method for 3D morphometry of human brain MRI. Because SynthSR works out of the box, combining it with existing tools is a cost- and time-efficient alternative to designing algorithms for a

specific type of MRI acquisition (orientation, resolution, and contrast). Such an approach not only would require human resources and specific anatomical and ML expertise to compile labeled datasets and train CNNs but also would be impractical when analyzing heterogeneous datasets. We note that, while SynthSR is compatible with any MRI acquisition, the quality of the output is higher for inputs with HR and contrast. SynthSR is publicly available as part of our widespread open-source neuroimaging software FreeSurfer, has no tunable parameters, and runs out of the box in a few seconds, making it very easy for anybody to use.

We have shown that SynthSR can be used in combination with the segmentation modules in FreeSurfer or FSL to chart trajectories of ROI volumes in aging and distinguish AD from control brains using a highly heterogeneous dataset of clinical scans, and we replicated the results of prior studies that used large numbers of research scans. This was made possible by the strong generalization ability provided by the DR strategy at training. Combined with a standard registration package like NiftyReg, SynthSR can also accurately register clinical data, including challenging cases, like scans of patients with glioma or stroke.

By facilitating analysis of brain MRI with existing tools, SynthSR has both clinical and research applications. In the clinic, SynthSR has the potential to enable 3D morphometry of anisotropic acquisitions (e.g., to measure longitudinal change, as in the BraTS-Reg experiment). In research, it has the potential to examine a variety of questions and may greatly increase the statistical power of many types of studies. For example, in studies of early disease stages, the increased sample sizes may enable the detection of subtle effects. Another example is the application to genome-wide and brain-wide association studies, where the required sample sizes are well into the thousands (21). SynthSR may also be used to facilitate the inclusion of minorities, which are often underrepresented in prospective studies but are more frequently scanned for clinical purposes.

Table 3. Volume of subcortical ROIs (ipsilateral and contralateral) in stroke. This table shows the mean volume of several representative ROIs in the contralateral and ipsilateral side of the stroke, the effect size between the two distributions (i.e., the asymmetry), and the correlation between the stroke lesion size and the volumes of the ROIs, including *P* values computed with Student's *t* test. We only included in this analysis stroke lesions that had more than two-thirds of their volume in one of the two hemispheres (*N* = 598).

Region	Hippocampus	Amygdala	Thalamus	Putamen	Caudate
Contralateral volume, mean and SD (mm ³)	4861 ± 560	2113 ± 274	7131 ± 803	5442 ± 686	4551 ± 483
Ipsilateral volume, mean and SD (mm ³)	4831 ± 560	1999 ± 310	5958 ± 1420	4642 ± 1147	4096 ± 655
Effect size: Contralateral versus ipsilateral	0.05	0.39	1.02	0.85	0.79
Correlation of lesion and contralateral volumes and <i>P</i> value	0.011 (<i>P</i> = 0.78)	−0.027 (<i>P</i> = 0.50)	0.021 (<i>P</i> = 0.61)	0.052 (<i>P</i> = 0.20)	0.07 (<i>P</i> = 0.07)
Correlation of lesion and ipsilateral volumes and <i>P</i> value	0.003 (<i>P</i> = 0.93)	−0.300 (<i>P</i> < 10 ^{−13})	−0.410 (<i>P</i> < 10 ^{−24})	−0.494 (<i>P</i> < 10 ^{−37})	−0.349 (<i>P</i> < 10 ^{−17})

SynthSR promises to increase the reproducibility of neuroimaging studies in two different manners: first, by being publicly available and not requiring any training or fine tuning; second, because it will deliver an increase in sample sizes that will, in turn, increase the reproducibility of studies. We emphasize that the purpose of SynthSR is not to produce images for clinical diagnosis but rather to generate synthetic scans with inpainted abnormalities that any neuroimaging software for 3D morphometry can use.

SynthSR has two main limitations. The first one is the fact that every tissue type (including pathological tissue) requires a label in the model. This is problematic when modeling continuous processes, e.g., tumor infiltration, which creates continuous gradients in the image intensities (due to partial voluming) that SynthSR does not model. Examples of this phenomenon are shown in fig. S2 (A to D). In future work, we plan to address this issue by compiling a dataset of lesions segmented at a finer level of detail and using it to retrain SynthSR, and also by using more complex models of pathology. The second limitation is that SynthSR is not guaranteed to yield plausible images, which can be problematic when inpainting large lesions. This is illustrated in fig. S2 (E and F), where SynthSR not only successfully inpaints a large stroke with white matter and cortical tissue but also hallucinates a smooth white matter boundary as well as gray matter in the contiguous inferior lateral ventricle. These problems could potentially be circumvented with additional GAN losses or by trying to quantify uncertainty.

In addition to tackling these two problems, future work will also explore more complex deep learning architectures. While we used a U-net (75, 76) that provides state-of-the-art results in many medical image segmentation and regression problems, the quick progress in architectures driven by the ML community will likely enable us to improve SynthSR with more sophisticated architectures in the coming years (e.g., transformers). Because millions of clinical brain MRI scans are acquired worldwide every year, we believe that tools that make them usable for quantitative analysis like SynthSR have the potential to transform human neuroimaging research and clinical neuroradiology.

MATERIALS AND METHODS

SynthSR relies on a CNN that requires a dataset of 1-mm isotropic 3D MPAGE scans and companion neuroanatomical labels for training. Then, the trained CNN can be used to analyze any other

dataset. In this section, we first describe the datasets used in training and in the experiments leading to the results reported above. Then, we describe the methods to train SynthSR and apply it to new data.

Datasets

Training dataset

Our approach, SynthSR, is trained on a subset of twenty 1-mm isotropic 3D MPAGE scans from the Open Access Series of Imaging Studies (OASIS) dataset (77), with dense neuroanatomical labels for cerebral and extracerebral ROIs. The cerebral ROIs comprise 36 brain regions and were manually labeled by members of our laboratory in previous work (8). The extracerebral regions were automatically obtained with a Bayesian approach using the atlas presented in (78), which includes labels for the eyes, skull, and soft tissue. The 20 subjects include 14 healthy volunteers and 6 subjects with probable AD, aged 51.3 ± 27.5 years (range, 18 to 82).

Test datasets

The experiments above rely on five different datasets. The first dataset comprises 9146 brain MRI scans from 1110 sessions, downloaded from the PACS at MGH under Institutional Review Board approval. The sessions correspond to subjects who had memory complaints and are not expected to have big lesions due to, e.g., tumors or strokes. This dataset is uncurated and includes a wide array of MRI modalities, including structural, angiography, and diffusion, among others. Virtually, every session is different from all others and includes different numbers of scans with different MRI modalities, contrasts, resolution, orientation, slice spacing, and thickness. Scans with more than three dimensions (typically diffusion) or with ICV under 1.1 liters (typically with partial field of view) were left out of the dataset. We refer to this dataset as the MGH dataset. In addition, we used four large public datasets that cover three different types of pathology: ADNI (5), ATLAS (69), and BraTS/BraTS-Reg (60, 63), which include subjects with AD, stroke, and brain tumors, respectively. Further details on the acquisitions can be found in the corresponding publications.

Data generation

Related work

SynthSR builds on our prior work on joint SR and synthesis using a combination of real and synthetic MRI scans (44). For the sake of completeness, we summarize our previous approach here. The method used a dataset of real 3D MPAGE scans (with normalized,

bias field–corrected intensities) and companion dense segmentations (including extracerebral tissue) to train a regression 3D CNN—a U-net (75, 76)—as follows. At every iteration, a random scan and corresponding segmentation were selected and geometrically augmented with a combination of random linear and nonlinear deformations (the latter, diffeomorphic). The deformed segmentation was then used to generate a synthetic scan with a generative model inspired by that of Bayesian segmentation, i.e., a Gaussian mixture model (GMM) conditioned on the underlying labels, combined with a model of additive noise and multiplicative bias field (10). Specific orientation, slice spacing, and slice thickness were then simulated using a model of partial voluming and subsampling (79). The CNN weights were then updated using this LR synthetic scan as input, and the 1-mm isotropic MPRAGE scan as ground-truth output. The GMM and resolution parameters were set to mimic the appearance of an arbitrarily prescribed type of acquisition, possibly multimodal (e.g., 4-mm coronal T2 plus 5-mm axial FLAIR).

As mentioned in Introduction, our newly proposed method addresses the three limitations of our previous approach, because of three key improvements (Fig. 7)—two of them related to data generation (discussed here), and a third related to the CNN architecture (in the next subsection).

Domain randomization

In our previous work, the synthetic scans used as input during training sought to mimic a specific type of acquisition. While this strategy enabled training of CNNs without specific training data (e.g.,

one could easily change the simulated spacing to 4 mm), it produced CNNs that were tailored to a specific acquisition. Here, we instead adopt a DR approach. DR has recently shown great success in robotics and computer vision (41). DR relies on training on a huge variety of simulations with fully randomized parameters—in our case, MR contrast (means and variances of ROIs), orientation (sagittal, axial, and coronal), slice spacing, slice thickness, bias field, and noise level (Fig. 7, A to C). By exposing the CNN to a different contrast, orientation, resolution, and noise level at every iteration during training, it learns not to rely on such features when estimating the output, thus becoming agnostic to them. Moreover, DR effectively addresses catastrophic forgetting (i.e., the fact that learning a new task may degrade performance in older task) (80) such that continual learning methods (81) are not required. The practical implications of DR on SynthSR are huge, since it bypasses the need to retrain and enables processing of heterogeneous datasets with a single CNN—all with a single FreeSurfer command:

```
mri_synthsr --i [input_scan] --o [synthetic_1mm_isotropic_t1]
```

Robustness to pathology

Modeling of pathology has two main aspects: shape and appearance—both of lesions and normal tissue. SynthSR seeks to preserve changes in shape of ROIs due to disease (e.g., atrophy) such that they can be captured by subsequent analyses to study clinical populations, e.g., as in the AD volumetry experiment above. This is achieved with a combination of two strategies: using a diverse training dataset including subjects with a wide age range (18 to 82 years)

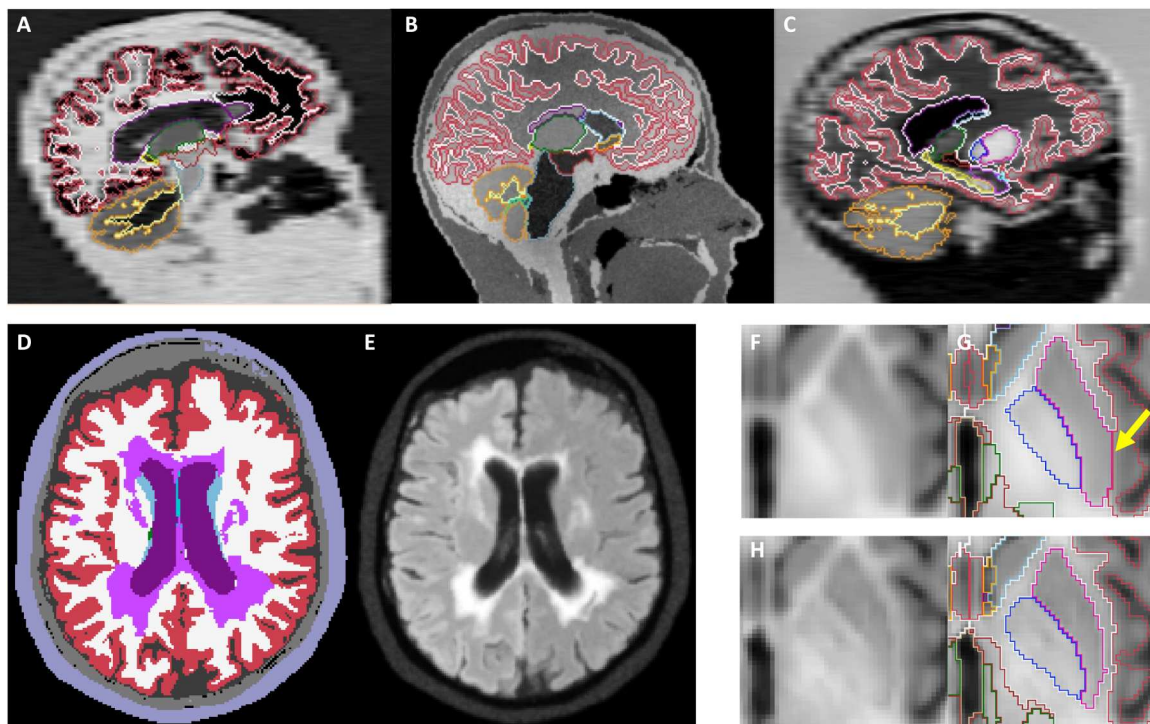


Fig. 7. Key improvements in SynthSR. Domain randomization: Sagittal slices of three synthetic samples with different orientation, contrast, and resolution are shown, along with the segmentations that were used to simulate them (A to C). Lesion modeling: Label maps augmented with lesion labels (D) are used to generate images with pathology (E). Auxiliary segmentation CNN: Blurry features are often obtained when minimizing only the error in estimated intensities (F), leading to segmentation errors (G). Encouraging the output image to be more accurately segmentable by a previously trained CNN noticeably sharpens the output (H) and ultimately leads to better segmentation (I).

and pathological atrophy (Alzheimer's), and aggressive shape augmentation of the 3D segmentations during training with random diffeomorphic deformations.

In terms of appearance, we have trained SynthSR to inpaint pathology with normal tissue to enable subsequent analysis with any morphometry tool. This is crucial when using methods that falter in the presence of lesions. Examples include the segmentation or registration of scans with strokes or tumors (as in our experiments with BraTS-Reg and ATLAS) or neuroimaging studies of multiple sclerosis, where it is common to segment white matter lesions and fill them with synthetic intensities resembling healthy white matter before processing with FreeSurfer or SPM (82, 83). To achieve this effect, we train the CNN with inputs where we randomly synthesize pathology, and outputs where no pathology is present. This is achieved as follows. First, we enhance the label maps of the training dataset with lesion segmentations registered from 60 held-out cases from ADNI, BRATS, and ATLAS (Fig. 7D). These new lesion labels are just treated as additional ROIs during image generation, creating synthetic lesions of random appearance during training (Fig. 7E). Second, we ensure that the regression targets (i.e., the MPRAGE images of the training dataset) do not display any pathology. Because these scans show some white matter lesions, we segment them with an open-source Bayesian method (84) and inpaint them with a publicly available algorithm (83) to obtain "clean" scans.

Generalization ability

The proposed method is extremely robust to overfitting. In terms of image intensities, SynthSR is practically immune to this phenomenon because it is trained with synthetic images generated with a model whose parameters (means, variances, and bias field) are fully randomized at every iteration during training. In terms of shape, overfitting is prevented by the aggressive geometric augmentation scheme described above, which includes strong random non-linear deformation of the 3D segmentations and random registration of random lesions at every iteration. Therefore, images of any orientation, resolution, and contrast can be fed to SynthSR without any preprocessing (e.g., no skull stripping or bias field correction is required), other than a simple minimum-maximum normalization of the intensities to comply with the range of inputs expected by the CNN.

CNN architecture

Our CNN builds on the architecture of our previous work, which is essentially a regression 3D U-net (75, 76). The U-net architecture is widespread in medical image segmentation and produces state-of-the-art accuracy in a wide array of tasks (85). Our previous work used a single CNN to regress the MPRAGE intensities from the input. However, this strategy often led to errors that, albeit small in terms of loss, could lead to large mistakes in subsequent tasks. For example, small errors in the predicted intensities may not influence the segmentation in the middle of the large structures, but can easily lead to mistakes around structures with faint boundaries (Fig. 7, F and G).

Here, we add a second CNN to the architecture: a segmentation U-net that is concatenated to the output of the regression U-net. This segmentation U-net is trained in advance, in a supervised fashion, using the training dataset (i.e., it predicts segmentations from MPRAGE scans). This CNN is frozen during training of SynthSR such that the regression network is encouraged to

synthesize images that can be accurately segmented by the supervised network, thus solving many of the problems of training only with regression (Fig. 7, H and I).

The regression and segmentation CNN have the same architecture, except for the final layer: The regression CNN has a linear layer with one feature, whereas the segmentation CNN has as many features as labels to segment, which are turned into probabilities with a softmax function. Both networks are 3D U-nets with five resolution levels, each consisting of two layers; each layer comprises a convolution with a $3 \times 3 \times 3$ kernel and a nonlinear exponential linear unit (ELU) activation. The first layer has 24 features. The number of features is doubled after each max pooling and halved after each up-sampling. The two U-nets are concatenated with the synthetic data generator into a single model completely implemented on the GPU for fast training, using Keras/TensorFlow.

Learning and inference

Training of our network minimizes a linear combination of two losses: a regression loss and a segmentation loss, combined with a relative weight λ . The goal is to minimize this loss with respect to the neural network parameters θ

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \mathcal{L}(\theta; X, Y, L_y) = \mathcal{L}_{\text{reg}}(\theta; X, Y) + \lambda \mathcal{L}_{\text{seg}}(\theta; X, L_y)$$

where $\hat{\theta}$ represents the optimal parameters, X is the input image, Y is the normalized ground-truth MPRAGE (with mean white matter intensity matched to 1.0), and L_y is the ground-truth segmentation of the MPRAGE. The regression loss is given by the L1 norm, i.e., the sum of absolute differences of the ground truth and predicted intensities

$$\mathcal{L}_{\text{reg}}(\theta; X, Y) = \|Y - f(X; \theta)\|_1$$

where $f(X; \theta)$ is the prediction for input image X when the CNN weights are equal to θ . The L1 norm produces crisper results than its L2 counterpart, i.e., sum of squared differences (44).

The segmentation loss is given by the (negated) soft average Dice score (overlap) between the ground-truth segmentation and the labels predicted by the segmentation U-net when given the predicted intensities (86)

$$\mathcal{L}_{\text{seg}} = 1 - \text{DICE}[L_y, \text{SEG}(f(X; \theta))]$$

where $\text{SEG}(\cdot)$ represents processing with the segmentation network. We emphasize that the use of these direct losses is much less likely to "hallucinate" features than those based on GANs, which produce highly realistic details that may not always be real (87).

This training loss is minimized using the Adam optimizer (88) with a learning rate of 10^{-4} and 250,000 iterations (approximately 2 weeks on an RTX 8000 GPU), after which the loss has converged by all practical measures. We set the relative weight to $\lambda = 0.25$ by visual inspection of the results on a small held-out dataset. The synthetic images are minimum-maximum normalized to the $[0,1]$ interval and up-scaled to 1-mm isotropic resolution such that the input and output images live in the same voxel space.

After training, one simply strips the synthetic generator and segmentation CNN from the model, as they are not needed for inference at test time. Input volumes are up-scaled to a resolution of $1 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$, padded to the closest multiple of 32 voxels in

each of the three spatial dimensions (required by the five resolution levels), minimum-maximum normalized to [0, 1], and processed with the regression CNN to estimate the synthetic MPRAGE intensities at 1-mm isotropic resolution. Inference takes less than 3 s on an RTX 8000 GPU and approximately 20 s on a modern CPU.

Statistical methods

We used Shapiro-Wilk tests (53) to assess whether data samples followed Gaussian distributions. Because this was never the case in our experiments, we used nonparametric statistics to compare populations: Wilcoxon rank sum tests to compare the medians (54) instead of *t* tests, and AUROCs instead of effect sizes.

Supplementary Materials

This PDF file includes:

Figs. S1 and S2

Other Supplementary Material for this manuscript includes the following:

Data file S1

REFERENCES AND NOTES

- B. Fischl, *FreeSurfer*. *Neuroimage* **62**, 774–781 (2012).
- M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, S. M. Smith, *FSL*. *Neuroimage* **62**, 782–790 (2012).
- W. D. Penny, K. J. Friston, J. T. Ashburner, S. J. Kiebel, T. E. Nichols, *Statistical Parametric Mapping: The Analysis of Functional Brain Images* (Elsevier, 2011).
- R. W. Cox, AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* **29**, 162–173 (1996).
- C. R. Jack Jr., M. A. Bernstein, N. C. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, P. J. Britson, J. L. Whitwell, C. Ward, A. M. Dale, J. P. Felmlee, J. L. Gunter, D. L. G. Hill, R. Killiany, N. Schuff, S. Fox-Bosetti, C. Lin, C. Studholme, C. S. De Carli, G. Krueger, H. A. Ward, G. J. Metzger, K. T. Scott, R. Mallozzi, D. Blezek, J. Levy, J. P. Debbs, A. S. Fleisher, M. Albert, R. Green, G. Bartzokis, G. Glover, J. Mugler, M. W. Weiner, The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging* **27**, 685–691 (2008).
- D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil; WU-Minn HCP Consortium, The WU-Minn human connectome project: An overview. *Neuroimage* **80**, 62–79 (2013).
- F. Alfaro-Almagro, M. Jenkinson, N. K. Bangerter, J. L. R. Andersson, L. Griffanti, G. Douaud, S. N. Sotiropoulos, S. Jbabdi, M. Hernandez-Fernandez, E. Vallee, D. Vidaurre, M. Webster, P. M. Carthy, C. Rorden, A. Daducci, D. C. Alexander, H. Zhang, I. Dragonu, P. M. Matthews, K. L. Miller, S. M. Smith, Image processing and quality control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage* **166**, 400–424 (2018).
- B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, A. M. Dale, Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron* **33**, 341–355 (2002).
- B. Patenaude, S. M. Smith, D. N. Kennedy, M. Jenkinson, A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage* **56**, 907–922 (2011).
- J. Ashburner, K. J. Friston, Unified segmentation. *Neuroimage* **26**, 839–851 (2005).
- D. N. Greve, B. Fischl, Accurate and robust brain image alignment using boundary-based registration. *Neuroimage* **48**, 63–72 (2009).
- J. L. Andersson, M. Jenkinson, S. Smith, Non-linear registration, aka Spatial normalisation (FMRIB Technical Report TR07JA2, FMRIB Analysis Group of the University of Oxford, 2007).
- M. Jenkinson, S. Smith, A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* **5**, 143–156 (2001).
- K. J. Friston, J. Ashburner, C. D. Frith, J.-B. Poline, J. D. Heather, R. S. J. Frackowiak, Spatial registration and normalization of images. *Hum. Brain Mapp.* **3**, 165–189 (1995).
- J. Ashburner, A fast diffeomorphic image registration algorithm. *Neuroimage* **38**, 95–113 (2007).
- Z. S. Saad, D. R. Glen, G. Chen, M. S. Beauchamp, R. Desai, R. W. Cox, A new method for improving functional-to-structural MRI alignment using local Pearson correlation. *Neuroimage* **44**, 839–848 (2009).
- J. P. Mugler III, J. R. Brookeman, Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP RAGE). *Magn. Reson. Med.* **15**, 152–157 (1990).
- J. Hennig, A. Nauerth, H. Friedburg, RARE imaging: A fast imaging method for clinical MR. *Magn. Reson. Med.* **3**, 823–833 (1986).
- O. Oren, E. Kebebew, J. P. A. Ioannidis, Curbing unnecessary and wasted diagnostic imaging. *JAMA* **321**, 245–246 (2019).
- P. M. Thompson, N. Jahanshad, C. R. K. Ching, L. E. Salminen, S. I. Thomopoulos, J. Bright, B. T. Baune, S. Bertolin, J. Bralten, W. B. Bruin, R. Bülow, J. Chen, Y. Chye, U. Dannlowski, C. G. F. de Kovel, G. Donohoe, L. T. Eyler, S. V. Faraone, P. Favre, C. A. Filippi, T. Frodl, D. Garijo, Y. Gil, H. J. Grabe, K. L. Grasby, T. Hajek, L. K. M. Han, S. N. Hatton, K. Hilbert, T. C. Ho, L. Holleran, G. Homuth, N. Hosten, J. Houenou, I. Ivanov, T. Jia, S. Kelly, M. Klein, J. S. Kwon, M. A. Laansma, J. Leerssen, U. Lueken, A. Nunes, J. O' Neill, N. Opel, F. Piras, F. Piras, M. C. Postema, E. Pozzi, N. Shatikhina, C. Soriano-Mas, G. Spalletta, D. Sun, A. Teumer, A. K. Tilot, L. Tozzi, C. van der Merwe, E. J. W. Van Someren, G. A. van Wingen, H. Völzke, E. Walton, L. Wang, A. M. Winkler, K. Wittfeld, M. J. Wright, J.-Y. Yun, G. Zhang, Y. Zhang-James, B. M. Adhikari, I. Agartz, M. Aghajani, A. Aleman, R. R. Althoff, A. Altmann, O. A. Andreassen, D. A. Baron, B. L. Bartnik-Olson, J. M. Bas-Hoogendam, A. R. Baskin-Sommers, C. E. Bearden, L. A. Berner, P. S. W. Boedhoe, R. M. Brouwer, J. K. Buitelaar, K. Caeyenberghs, C. A. M. Cecil, R. A. Cohen, J. H. Cole, P. J. Conrod, S. A. De Brito, S. M. C. de Zwart, E. L. Dennis, S. Desrivieres, D. Dima, S. Ehrlich, C. Esopenko, G. Fairchild, S. E. Fisher, J.-P. Fouche, C. Francks, S. Frangou, B. Franke, H. P. Garavan, D. C. Glahn, N. A. Groenewold, T. P. Gurbolt, B. A. Gutman, T. Hahn, I. H. Harding, D. Hernaus, D. P. Hibar, F. G. Hillary, M. Hoogman, H. E. Hulshoff Pol, M. Jalbrzikowski, G. A. Karkashadze, E. T. Klapwijk, R. C. Knickmeyer, P. Kochunov, I. K. Koerte, X.-Z. Kong, S.-L. Liew, A. P. Lin, M. W. Logue, E. Luders, F. Maciardi, S. Mackey, A. R. Mayer, C. R. Mc Donald, A. B. Mc Mahon, S. E. Medland, G. Modinos, R. A. Morey, S. C. Mueller, P. Mukherjee, L. Namazova-Baranova, T. M. Nir, A. Olsen, P. Paschou, D. S. Pine, F. Pizzagalli, M. E. Rentería, J. D. Rohrer, P. G. Sämann, L. Schmaal, G. Schumann, M. S. Shiroishi, S. M. Sisodiya, D. J. A. Smit, I. E. Sønderby, D. J. Stein, J. L. Stein, M. Tahmasian, D. F. Tate, J. A. Turner, O. A. van den Heuvel, N. J. A. van der Wee, Y. D. van der Werf, T. G. M. van Erp, N. E. M. van Haren, D. van Rooij, L. S. van Velden, I. M. Veer, D. J. Veltman, J. E. Villalon-Reina, H. Walter, C. D. Whelan, E. A. Wilde, M. Zarei, V. Zelman; for the ENIGMA Consortium, ENIGMA and global neuroscience: A decade of large-scale studies of the brain in health and disease across more than 40 countries. *Transl. Psychiatry* **10**, 100 (2020).
- S. Marek, B. Tervo-Clemmens, F. J. Calabro, D. F. Montez, B. P. Kay, A. S. Hatoum, M. R. Donohue, W. Foran, R. L. Miller, T. J. Hendrickson, S. M. Malone, S. Kandala, E. Feczko, O. Miranda-Dominguez, A. M. Graham, E. A. Earl, A. J. Perrone, M. Cordova, O. Doyle, L. A. Moore, G. M. Conan, J. Uriarte, K. Snider, B. J. Lynch, J. C. Wilgenbusch, T. Pengo, A. Tam, J. Chen, D. J. Newbold, A. Zheng, N. A. Seider, A. N. Van, A. Metoki, R. J. Chauvin, T. O. Laumann, D. J. Greene, S. E. Petersen, H. Garavan, W. K. Thompson, T. E. Nichols, B. T. T. Yeo, D. M. Barch, B. Luna, D. A. Fair, N. U. F. Dosenbach, Reproducible brain-wide association studies require thousands of individuals. *Nature* **603**, 654–660 (2022).
- Z. Wang, J. Chen, S. C. Hoi, Deep learning for image super-resolution: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **43**, 3365–3387 (2020).
- W. Yang, X. Zhang, Y. Tian, W. Wang, J.-H. Xue, Q. Liao, Deep learning for single image super-resolution: A brief review. *IEEE Trans. Multimedia* **21**, 3106–3121 (2019).
- C.-H. Pham, C. Tor-Díez, H. Meunier, N. Bednarek, R. Fablet, N. Passat, F. Rousseau, Multi-scale brain MRI super-resolution using deep 3D convolutional networks. *Comput. Med. Imaging Graph.* **77**, 101647 (2019).
- R. Tanno, D. E. Worrall, A. Ghosh, E. Kaden, S. N. Sotiropoulos, A. Criminisi, D. C. Alexander, in *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention* (Springer, 2017), pp. 611–619.
- Q. Tian, B. Bilgic, Q. Fan, C. Ngamsombat, N. Zaretskaya, N. E. Fultz, N. A. Ohringer, A. S. Chaudhari, Y. Hu, T. Witzel, K. Setsompop, J. R. Polimeni, S. Y. Huang, Improving *in vivo* human cerebral cortical surface reconstruction using data-driven super-resolution. *Cereb. Cortex* **31**, 463–482 (2021).
- I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets. *Adv. Neural. Inf. Process. Syst.* **27** (2014).
- Y. Chen, F. Shi, A. G. Christodoulou, Y. Xie, Z. Zhou, D. Li, in *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention* (Springer, 2018), pp. 91–99.
- Q. Lyu, H. Shan, C. Steber, C. Helis, C. Whitlow, M. Chan, G. Wang, Multi-contrast super-resolution MRI through a progressive network. *IEEE Trans. Med. Imaging* **39**, 2738–2749 (2020).
- Q. Delannoy, C.-H. Pham, C. Cazorla, C. Tor-Díez, G. Dollé, H. Meunier, N. Bednarek, R. Fablet, N. Passat, F. Rousseau, SegSRGAN: Super-resolution and segmentation using generative adversarial networks—Application to neonatal brain MRI. *Comput. Biol. Med.* **120**, 103755 (2020).
- J.-Y. Zhu, T. Park, P. Isola, A. A. Efros, in *Proceedings of the IEEE International Conference on Computer Vision* (IEEE, 2017), pp. 2223–2232.

32. T.-A. Song, S. R. Chowdhury, F. Yang, J. Dutta, PET image super-resolution using generative adversarial networks. *Neural Netw.* **125**, 83–91 (2020).
33. A. Chartsias, T. Joyce, M. V. Giuffrida, S. A. Tsafaris, Multimodal MR synthesis via modality-invariant latent representation. *IEEE Trans. Med. Imaging* **37**, 803–814 (2018).
34. L. Xiang, Q. Wang, D. Nie, L. Zhang, X. Jin, Y. Qiao, D. Shen, Deep embedding convolutional neural network for synthesizing CT image from T1-weighted MR image. *Med. Image Anal.* **47**, 31–44 (2018).
35. D. Nie, R. Trullo, J. Lian, L. Wang, C. Petitjean, S. Ruan, Q. Wang, D. Shen, Medical image synthesis with deep convolutional adversarial networks. *IEEE Trans. Biomed. Eng.* **65**, 2720–2730 (2018).
36. S. U. Dar, M. Yurt, L. Karacan, A. Erdem, E. Erdem, T. Çukur, Image synthesis in multi-contrast MRI with conditional generative adversarial networks. *IEEE Trans. Med. Imaging* **38**, 2375–2388 (2019).
37. Y. Fei, B. Zhan, M. Hong, X. Wu, J. Zhou, Y. Wang, Deep learning-based multi-modal computing with feature disentanglement for MRI image synthesis. *Med. Phys.* **48**, 3778–3789 (2021).
38. G. M. Conte, A. D. Weston, D. C. Vogelsang, K. A. Philbrick, J. C. Cai, M. Barbera, F. Sanvito, D. H. Lachance, R. B. Jenkins, W. O. Tobin, J. E. Eckel-Passow, B. J. Erickson, Generative adversarial networks to synthesize missing T1 and FLAIR MRI sequences for use in a multisequence brain tumor segmentation model. *Radiology* **299**, 313–323 (2021).
39. L. Qu, Y. Zhang, S. Wang, P.-T. Yap, D. Shen, Synthesized 7T MRI from 3T MRI via deep learning in spatial and wavelet domains. *Med. Image Anal.* **62**, 101663 (2020).
40. B. Billot, D. N. Greve, K. Van Leemput, B. Fischl, J. E. Iglesias, A. V. Dalca, A learning strategy for contrast-agnostic MRI segmentation. *Medical Imaging Deep Learn.* **121**, 75–93 (2020).
41. C. Shorten, T. M. Khoshgoftaar, A survey on image data augmentation for deep learning. *J. Big Data* **6**, 60 (2019).
42. M. Wang, W. Deng, Deep visual domain adaptation: A survey. *Neurocomputing* **312**, 135–153 (2018).
43. C. Zhao, B. E. Dewey, D. L. Pham, P. A. Calabresi, D. S. Reich, J. L. Prince, SMORE: A self-supervised anti-aliasing and super-resolution algorithm for MRI using deep learning. *IEEE Trans. Med. Imaging* **40**, 805–817 (2021).
44. J. E. Iglesias, B. Billot, Y. Balbastre, A. Tabari, J. Conklin, R. G. González, D. C. Alexander, P. Golland, B. L. Edlow, B. Fischl; Alzheimer's Disease Neuroimaging Initiative, Joint super-resolution and synthesis of 1 mm isotropic MP-RAGE volumes from clinical MRI exams with scans of different orientation, resolution and contrast. *Neuroimage* **237**, 118206 (2021).
45. J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, P. Abbeel, in *Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2017), pp. 23–30.
46. C. Guo, D. Ferreira, K. Fink, E. Westman, T. Granberg, Repeatability and reproducibility of FreeSurfer, FSL-SIENAX and SPM brain volumetric measurements and the effect of lesion filling in multiple sclerosis. *Eur. Radiol.* **29**, 1355–1364 (2019).
47. R. A. I. Bethlehem, J. Seidlitz, S. R. White, J. W. Vogel, K. M. Anderson, C. Adamson, S. Adler, G. S. Alexopoulos, E. Anagnostou, A. Areces-Gonzalez, D. E. Astle, B. Auyeung, M. Ayub, J. Bae, G. Ball, S. Baron-Cohen, R. Beare, S. A. Bedford, V. Benegal, F. Beyer, J. Blangero, M. B. Cabez, J. P. Boardman, M. Borzage, J. F. Bosch-Bayard, N. Bourke, V. D. Calhoun, M. M. Chakravarty, C. Chen, C. Chertavian, G. Chetelat, Y. S. Chong, J. H. Cole, A. Corvin, M. Costantino, E. Courchesne, F. Crivello, V. L. Cropley, J. Crosbie, N. Crossley, M. Delarue, R. Delorme, S. Desrivieres, G. A. Devenyi, M. A. Di Biase, R. Dolan, K. A. Donald, G. Donohoe, K. Dunlop, A. D. Edwards, J. T. Ellison, C. T. Ellis, J. A. Elman, L. Eyster, D. A. Fair, E. Feczko, P. C. Fletcher, P. Fonagy, C. E. Franz, L. Galan-Garcia, A. Gholipour, J. Giedd, J. H. Gilmore, D. C. Glahn, I. M. Goodyer, P. E. Grant, N. A. Groenewold, F. M. Gunning, R. E. Gur, R. C. Gur, C. F. Hammill, O. Hansson, T. Hedden, A. Heinz, R. N. Henson, K. Heuer, J. Hoare, B. Holla, A. J. Holmes, R. Holt, H. Huang, K. Im, J. Ipser, C. R. Jack Jr, A. P. Jackowski, T. Jia, K. A. Johnson, P. B. Jones, D. T. Jones, R. S. Kahn, H. Karlsson, L. Karlsson, R. Kawashima, E. A. Kelley, S. Kern, K. W. Kim, M. G. Kitzbichler, W. S. Kremen, F. Lalonde, B. Landeau, S. Lee, J. Lerch, J. D. Lewis, J. Li, W. Liao, C. Liston, M. V. Lombardo, J. Lv, C. Lynch, T. T. Mallard, M. Marcellis, R. D. Markello, S. R. Mathias, B. Mazoyer, P. McGuire, M. J. Meaney, A. Mechelli, N. Medic, B. Misic, S. E. Morgan, D. Mothersill, J. Nigg, M. Q. W. Ong, C. Ortinau, R. Ossenkoppele, M. Ouyang, L. Palaniyappan, L. Paly, P. M. Pan, C. Pantelis, M. M. Park, T. Paus, Z. Pausova, D. Paz-Linares, A. P. Binette, K. Pierce, X. Qian, J. Qiu, A. Qiu, A. Raznahan, T. Rittman, A. Rodrigue, C. K. Rollins, R. Romero-Garcia, L. Ronan, M. D. Rosenberg, D. H. Rowitch, G. A. Salum, T. D. Satterthwaite, H. L. Schaare, R. J. Schachar, A. P. Schultz, G. Schumann, M. Schöll, D. Sharp, R. T. Shinohara, I. Skoog, C. D. Smyser, R. A. Sperling, D. J. Stein, A. Stolicyn, J. Suckling, G. Sullivan, Y. Taki, B. Thyreau, R. Toro, N. Traut, K. A. Tsvetanov, N. B. Turk-Browne, J. J. Tuulari, C. Tzourio, É. Vachon-Presseau, M. J. Valdes-Sosa, P. A. Valdes-Sosa, S. L. Valk, T. van Amelsvoort, S. N. Vandekar, L. Vasung, L. W. Victoria, S. Villeneuve, A. Villringer, P. E. Vértes, K. Wagstyl, Y. S. Wang, S. K. Warfield, V. Warrier, E. Westman, M. L. Westwater, H. C. Whalley, A. V. Witte, N. Yang, B. Yeo, H. Yun, A. Zalesky, H. J. Zar, A. Zettergren, J. H. Zhou, H. Ziauddeen, A. Zugman, X. N. Zuo; 3R-BRAIN; AIBL; Alzheimer's Disease Neuroimaging Initiative; Alzheimer's Disease Repository Without Borders Investigators; CALM Team; Cam-CAN; CCNP; COBRE; cVEDA; ENIGMA
- Developmental Brain Age Working Group; Developing Human Connectome Project; FinnBrain; Harvard Aging Brain Study; IMAGEN; KNE96; The Mayo Clinic Study of Aging; NSPN; POND; The PREVENT-AD Research Group; VETSA; E. T. Bullmore, A. F. Alexander-Bloch, Brain charts for the human lifespan. *Nature* **604**, 525–533 (2022).
48. D. Dima, A. Modabbernia, E. Papachristou, G. E. Doucet, I. Agartz, M. Aghajani, T. N. Akudjedu, A. Albajes-Eizaguirre, D. Alnaes, K. I. Alpert, M. Andersson, N. C. Andreasen, O. A. Andreassen, P. Asherson, T. Banaschewski, N. Bargallo, S. Baumeister, R. Baur-Streubel, A. Bertolino, A. Bonvino, D. I. Boomsma, S. Borgwardt, J. Bourque, D. Brandeis, A. Breier, H. Brodaty, R. M. Brouwer, J. K. Buitelaar, G. F. Busatto, R. L. Buckner, V. Calhoun, E. J. Canales-Rodríguez, D. M. Cannon, X. Caseras, F. X. Castellanos, S. Cervenka, T. M. Chaim-Avancini, C. R. K. Ching, V. Chubar, V. P. Clark, P. Conrod, A. Conzelmann, B. Crespo-Facorro, F. Crivello, E. A. Crone, U. Dannlowski, A. M. Dale, C. Davey, Eco J C de Geus, L. de Haan, G. I. de Zubicaray, A. den Braber, E. W. Dickie, A. D. Giorgio, N. T. Doan, E. S. Dorum, S. Ehrlich, S. Erker, T. Espeseth, H. Fatouros-Bergman, S. E. Fisher, J.-P. Fouché, B. Franke, T. Frodl, P. Fuentes-Claramonte, D. C. Glahn, I. H. Gotlib, H.-J. Grabe, O. Grimm, N. A. Groenewold, D. Grotegerd, O. Gruber, P. Gruner, R. E. Gur, R. C. Gur, T. Hahn, B. J. Harrison, C. A. Hartman, S. N. Hatton, A. Heinz, D. J. Heslenfeld, D. P. Hibar, I. B. Hickie, B.-C. Ho, P. J. Hoekstra, S. Hohmann, A. J. Holmes, M. Hoogman, N. Hosten, F. M. Howells, H. E. H. Pol, C. Huysler, N. Jahanshad, A. James, T. L. Jerinigan, J. Jiang, E. G. Jönsson, J. A. Joska, R. Kahn, A. Kalnin, R. Kanai, M. Klein, T. P. Klyushnik, L. Koenders, S. Koops, B. Krämer, J. Kuntsi, J. Lagopoulos, L. Lázaro, I. Lebedeva, W. H. Lee, K.-P. Lesch, C. Lochner, M. W. J. Machielsen, S. Maingault, N. G. Martin, I. Martínez-Zalacain, D. Mataix-Cols, B. Mazoyer, C. M. Donald, B. C. McDonald, A. M. McIntosh, K. L. McMahon, G. M. Philemy, S. Meinert, J. M. Menchón, S. E. Medland, A. Meyer-Lindenberg, J. Naaijen, P. Najt, T. Nakao, J. E. Nordvik, L. Nyberg, J. Oosterlaan, V. O.-G. de la Foz, Y. Paloyelis, P. Pauli, G. Pergola, E. Pomarol-Clotet, M. J. Portella, S. G. Potkin, J. Radua, A. Reif, D. A. Rinker, J. L. Roffman, P. G. P. Rosa, M. D. Sacchet, P. S. Sachdev, R. Salvador, P. Sánchez-Juan, S. Sarró, T. D. Satterthwaite, A. J. Saykin, M. H. Serpa, L. Schmaal, K. Schnell, G. Schumann, K. Sim, J. W. Smoller, I. Sommer, C. Soriano-Mas, D. J. Stein, L. T. Strike, S. C. Swagerman, C. K. Tamnes, H. S. Temmingh, S. I. Thomopoulos, A. S. Tomyshev, D. Tordesillas-Gutiérrez, J. N. Trollor, J. A. Turner, A. Uhlmann, O. A. van den Heuvel, D. van den Meer, N. J. A. van der Wee, N. E. M. van Haren, D. Van't Ent, T. G. M. van Erp, I. M. Veer, D. J. Veltman, A. Voineskos, H. Völzke, H. Walter, E. Walton, L. Wang, Y. Wang, T. H. Wassink, B. Weber, W. Wen, J. D. West, L. T. Westlye, H. Whalley, L. M. Wierenga, S. C. R. Williams, K. Wittfeld, D. H. Wolf, A. Worker, M. J. Wright, K. Yang, Y. Yoncheva, M. V. Zanetti, G. C. Ziegler, P. M. Thompson, S. Franouk; Karolinska Schizophrenia Project (KaSP), Subcortical volumes across the lifespan: Data from 18,605 healthy individuals aged 3–90 years. *Hum. Brain Mapp.* **43**, 452–469 (2022).
49. P. Coupé, G. Catheline, E. Lanuza, J. V. Manjón; Alzheimer's Disease Neuroimaging Initiative, Towards a unified analysis of brain maturation and aging across the entire lifespan: A MRI analysis. *Hum. Brain Mapp.* **38**, 5501–5518 (2017).
50. A. Convit, M. De Leon, C. Tarshish, S. De Santi, W. Tsui, H. Rusinek, A. George, Specific hippocampal volume reductions in individuals at risk for Alzheimer's disease. *Neurobiol. Aging* **18**, 131–138 (1997).
51. K. M. Gosche, J. A. Mortimer, C. D. Smith, W. R. Markesbery, D. A. Snowdon, Hippocampal volume as an index of Alzheimer neuropathology: Findings from the Nun study. *Neurology* **58**, 1476–1482 (2002).
52. B. Dubois, H. H. Feldman, C. Jacova, S. T. Dekosky, P. Barberger-Gateau, J. Cummings, A. Delacourte, D. Galasko, S. Gauthier, G. Jicha, K. Meguro, J. O'Brien, F. Pasquier, P. Robert, M. Rossor, S. Salloway, Y. Stern, P. J. Visser, P. Scheltens, Research criteria for the diagnosis of Alzheimer's disease: Revising the NINCDS–ADRDA criteria. *Lancet Neurol.* **6**, 734–746 (2007).
53. S. S. Shapiro, M. B. Wilk, An analysis of variance test for normality (complete samples). *Biometrika* **52**, 591–611 (1965).
54. F. Wilcoxon, in *Breakthroughs in Statistics* (Springer, 1992), pp. 196–202.
55. A. Sotiras, C. Davatzikos, N. Paragios, Deformable medical image registration: A survey. *IEEE Trans. Med. Imaging* **32**, 1153–1190 (2013).
56. M. Reuter, N. J. Schmansky, H. D. Rosas, B. Fischl, Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage* **61**, 1402–1418 (2012).
57. R. P. Woods, J. C. Mazziotta, S. R. Cherry, MRI-PET registration with automated algorithm. *J. Comput. Assist. Tomogr.* **17**, 536–546 (1993).
58. A. C. Evans, D. L. Collins, S. Mills, E. D. Brown, R. L. Kelly, T. M. Peters, in *Proceedings of the 1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference* (IEEE, 1993), pp. 1813–1817.
59. H. Akbari, L. Macyszyn, X. Da, M. Bilello, R. L. Wolf, M. Martinez-Lage, G. Biros, M. Alonso-Basanta, D. M. O'Rourke, C. Davatzikos, Imaging surrogates of infiltration obtained via multiparametric imaging pattern analysis predict subsequent location of recurrence of glioblastoma. *Neurosurgery* **78**, 572–580 (2016).
60. B. Baheti, D. Waldmannstetter, S. Chakrabarty, H. Akbari, M. Bilello, B. Wiestler, J. Schwarting, E. Calabrese, J. Rudie, S. Abidi, The brain tumor sequence registration

- challenge: Establishing correspondence between pre-operative and follow-up MRI scans of diffuse glioma patients. *arXiv:2112.06979 [eess.IV]* (13 December 2021).
61. M. Modat, D. M. Cash, P. Daga, G. P. Winston, J. S. Duncan, S. Ourselin, Global image registration using a symmetric block-matching approach. *J. Med. Imaging* **1**, 024003 (2014).
 62. M. Modat, G. R. Ridgway, Z. A. Taylor, M. Lehmann, J. Barnes, D. J. Hawkes, N. C. Fox, S. Ourselin, Fast free-form deformation using graphics processing units. *Comput. Methods Programs Biomed.* **98**, 278–284 (2010).
 63. B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, L. Lanczi, E. Gerstner, M.-A. Weber, T. Arbel, B. B. Avants, N. Ayache, P. Buendia, D. L. Collins, N. Cordier, J. J. Corso, A. Criminisi, T. Das, H. Delingette, Ç. Demiralp, C. R. Durst, M. Dojat, S. Doyle, J. Festa, F. Forbes, E. Geremia, B. Glocker, P. Golland, X. Guo, A. Hamamci, K. M. Iftekharrudin, R. Jena, N. M. John, A. E. Konukoglu, D. Lashkari, J. A. Mariz, R. Meier, S. Pereira, D. Precup, S. J. Price, T. R. Raviv, S. M. S. Reza, M. Ryan, D. Sarikaya, L. Schwartz, H.-C. Shin, J. Shotton, C. A. Silva, N. Sousa, N. K. Subbanna, G. Szekely, T. J. Taylor, O. M. Thomas, N. J. Tustison, G. Unal, F. Vasseur, M. Wintermark, D. H. Ye, L. Zhao, B. Zhao, D. Zikic, M. Prastawa, M. Reyes, K. Van Leemput, The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans. Med. Imaging* **34**, 1993–2024 (2015).
 64. S. Joshi, B. Davis, M. Jomier, G. Gerig, Unbiased diffeomorphic atlas construction for computational anatomy. *Neuroimage* **23**, S151–S160 (2004).
 65. M. Bilello, H. Akbari, X. Da, J. M. Pisapia, S. Mohan, R. L. Wolf, D. M. O'Rourke, M. Martinez-Lage, C. Davatzikos, Population-based MRI atlases of spatial distribution are specific to patient and tumor characteristics in glioblastoma. *NeuroImage Clin.* **12**, 34–40 (2016).
 66. A. Zavaliangos-Petropulu, B. Lo, M. R. Donnelly, N. Schweighofer, K. Lohse, N. Jahanshad, G. Barisano, N. Banaj, M. R. Borich, L. A. Boyd, C. M. Buettefisch, W. D. Byblow, J. M. Cassidy, C. C. Charalambous, A. B. Conforto, J. A. Di Carlo, A. N. Dula, N. Egorova-Brumley, M. R. Etherton, W. Feng, K. A. Fercho, F. Geranmayeh, C. A. Hanlon, K. S. Hayward, B. Hordacre, S. A. Kautz, M. S. Khlif, H. Kim, A. Kuceyeski, D. J. Lin, J. Liu, M. Lotze, B. J. M. Intosh, J. L. Margetis, F. B. Mohamed, F. Piras, A. Ramos-Murguialday, K. P. Revill, P. S. Roberts, A. D. Robertson, H. M. Schambra, N. J. Seo, M. S. Shiroishi, C. M. Stinear, S. R. Soekadar, G. Spalletta, M. Taga, W. K. Tang, G. T. Thielman, D. Vecchio, N. S. Ward, L. T. Westlye, E. Werden, C. Winstein, G. F. Wittenberg, S. L. Wolf, K. A. Wong, C. Yu, A. Brodtmann, S. C. Cramer, P. M. Thompson, S.-L. Liew, Chronic stroke sensorimotor impairment is related to smaller hippocampal volumes: An ENIGMA analysis. *J. Am. Heart Assoc.* **11**, e025109 (2022).
 67. V. C. T. Mok, B. Y. K. Lam, A. Wong, H. Ko, H. S. Markus, L. K. S. Wong, Early-onset and delayed-onset poststroke dementia—Revisiting the mechanisms. *Nat. Rev. Neurol.* **13**, 148–159 (2017).
 68. S.-L. Liew, A. Zavaliangos-Petropulu, N. Schweighofer, N. Jahanshad, C. E. Lang, K. R. Lohse, N. Banaj, G. Barisano, L. A. Baugh, A. K. Bhattacharya, B. Bigjahan, M. R. Borich, L. A. Boyd, A. Brodtmann, C. M. Buettefisch, W. D. Byblow, J. M. Cassidy, C. C. Charalambous, V. Ciullo, A. B. Conforto, R. C. Craddock, A. N. Dula, N. Egorova, W. Feng, K. A. Fercho, C. M. Gregory, C. A. Hanlon, K. S. Hayward, J. A. Holguin, B. Hordacre, D. H. Hwang, S. A. Kautz, M. S. Khlif, B. Kim, H. Kim, A. Kuceyeski, B. Lo, J. Liu, D. Lin, M. Lotze, B. J. M. Intosh, J. L. Margetis, F. B. Mohamed, J. E. Nordvik, M. A. Petoe, F. Piras, S. Raju, A. Ramos-Murguialday, K. P. Revill, P. S. Roberts, A. D. Robertson, H. M. Schambra, N. J. Seo, M. S. Shiroishi, S. R. Soekadar, G. Spalletta, C. M. Stinear, A. Suri, W. K. Tang, G. T. Thielman, V. N. Thijs, D. Vecchio, N. S. Ward, L. T. Westlye, C. J. Winstein, G. F. Wittenberg, K. A. Wong, C. Yu, S. L. Wolf, S. C. Cramer, P. M. Thompson; ENIGMA Stroke Recovery Working Group, Smaller spared subcortical nuclei are associated with worse post-stroke sensorimotor outcomes in 28 cohorts worldwide. *Brain Commun.* **3**, fcb254 (2021).
 69. S.-L. Liew, J. M. Anglin, N. W. Banks, M. Sondag, K. L. Ito, H. Kim, J. Chan, J. Ito, C. Jung, N. Khoshab, S. Lefebvre, W. Nakamura, D. Saldana, A. Schmiesing, C. Tran, D. Vo, T. Ard, P. Heydari, B. Kim, L. Aziz-Zadeh, S. C. Cramer, J. Liu, S. Soekadar, J.-E. Nordvik, L. T. Westlye, J. Wang, C. Winstein, C. Yu, L. Ai, B. Koo, R. C. Craddock, M. Milham, M. Lakich, A. Pienta, A. Stroud, A large, open source dataset of stroke anatomical brain images and manual lesion segmentations. *Sci. Data* **5**, 180011 (2018).
 70. S.-L. Liew, B. P. Lo, M. R. Donnelly, A. Zavaliangos-Petropulu, J. N. Jeong, G. Barisano, A. Hutton, J. P. Simon, J. M. Juliano, A. Suri, Z. Wang, A. Abdullah, J. Kim, T. Ard, N. Banaj, M. R. Borich, L. A. Boyd, A. Brodtmann, C. M. Buettefisch, L. Cao, J. M. Cassidy, V. Ciullo, A. B. Conforto, S. C. Cramer, R. Dacosta-Aguayo, E. de la Rosa, M. Domin, A. N. Dula, W. Feng, A. R. Franco, F. Geranmayeh, A. Gramfort, C. M. Gregory, C. A. Hanlon, B. G. Hordacre, S. A. Kautz, M. S. Khlif, H. Kim, J. S. Kirschke, J. Liu, M. Lotze, B. J. MacIntosh, M. Mataró, F. B. Mohamed, J. E. Nordvik, G. Park, A. Pienta, F. Piras, S. M. Redman, K. P. Revill, M. Reyes, A. D. Robertson, N. J. Seo, S. R. Soekadar, G. Spalletta, A. Sweet, M. Telenczuk, G. Thielman, L. T. Westlye, C. J. Winstein, G. F. Wittenberg, K. A. Wong, C. Yu, A large, curated, open-source stroke neuroimaging dataset to improve lesion segmentation algorithms. *Sci. Data* **9**, 320 (2022).
 71. M. S. Khlif, E. Werden, N. Egorova, M. Boccardi, A. Redolfi, L. Bird, A. Brodtmann, Assessment of longitudinal hippocampal atrophy in the first year after ischemic stroke using automatic segmentation techniques. *NeuroImage Clin.* **24**, 102008 (2019).
 72. L. C. Krishnamurthy, G. N. Champion, K. M. McGregor, V. Krishnamurthy, A. Turabi, S. R. Roberts, J. R. Nocera, M. R. Borich, A. D. Rodriguez, S. R. Belagaje, R. M. Harrington, M. L. Harris-Love, S. M. Harnish, J. H. Drucker, M. Benjamin, M. L. Meadows, L. Seeds, Z. Z. Zlatar, A. Sudhyadhom, A. J. Butler, A. Garcia, C. Patten, J. Trinastic, S. A. Kautz, C. Gregory, B. A. Crosson, The effect of time since stroke, gender, age, and lesion size on thalamus volume in chronic stroke: A pilot study. *Sci. Rep.* **10**, 20488 (2020).
 73. J. C. L. Looi, V. Tatham, R. Kumar, J. J. Maller, E. Millard, W. Wen, X. Chen, H. Brodaty, P. Sachdev, Caudate nucleus volumes in stroke and vascular dementia. *Psychiatry Res.* **174**, 67–75 (2009).
 74. S. Koch, S. Mueller, M. Foddis, T. Bienert, D. von Elverfeldt, F. Knab, T. D. Farr, R. Bernard, M. Dopatka, A. Rex, U. Dirnagl, C. Harms, P. Boehm-Sturm, Atlas registration for edema-corrected MRI lesion volume in mouse stroke models. *J. Cereb. Blood Flow Metab.* **39**, 313–323 (2019).
 75. Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, O. Ronneberger, in *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention* (Springer, 2016), pp. 424–432.
 76. O. Ronneberger, P. Fischer, T. Brox, in *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention* (Springer, 2015), pp. 234–241.
 77. D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, R. L. Buckner, Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *J. Cogn. Neurosci.* **19**, 1498–1507 (2007).
 78. O. Puonti, J. E. Iglesias, K. Van Leemput, Fast and sequence-adaptive whole-brain segmentation using parametric Bayesian modeling. *Neuroimage* **143**, 235–249 (2016).
 79. K. Van Leemput, F. Maes, D. Vandermeulen, P. Suetens, A unifying framework for partial volume segmentation of brain MR images. *IEEE Trans. Med. Imaging* **22**, 105–119 (2003).
 80. R. Volpi, D. Larlus, G. Rogez, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (IEEE, 2021), pp. 4443–4453.
 81. M. Delange, R. Aljundi, M. Masana, S. Parisot, X. Jia, A. Leonardis, G. Slabaugh, T. Tuytelaars, A continual learning survey: Defying forgetting in classification tasks. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**, 3366–3385 (2021).
 82. D. Pareto, J. Sastre-Garriga, F. Aymerich, C. Auger, M. Tintoré, X. Montalban, A. Rovira, Lesion filling effect in regional brain volume estimations: A study in multiple sclerosis patients with low lesion load. *Neuroradiology* **58**, 467–474 (2016).
 83. F. Prados, M. J. Cardoso, B. Kanber, O. Ciccarelli, R. Kapoor, C. A. G. Wheeler-Kingshott, S. Ourselin, A multi-time-point modality-agnostic patch-based method for lesion filling in multiple sclerosis. *Neuroimage* **139**, 376–384 (2016).
 84. S. Cerri, O. Puonti, D. S. Meier, J. Wuerfel, M. Mühlau, H. R. Siebner, K. Van Leemput, A contrast-adaptive method for simultaneous whole-brain and lesion segmentation in multiple sclerosis. *Neuroimage* **225**, 117471 (2021).
 85. F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, K. H. Maier-Hein, nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**, 203–211 (2021).
 86. F. Milletari, N. Navab, S.-A. Ahmadi, in *Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV)* (IEEE, 2016), pp. 565–571.
 87. J. P. Cohen, M. Luck, S. Honari, in *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention* (Springer, 2018), pp. 529–536.
 88. D. P. Kingma, J. Ba, Adam: A method for stochastic optimization. *arXiv:1412.6980 [cs.LG]* (22 December 2014).

Acknowledgments: We thank W. T. Kimberly and K. Sheth for consultation regarding the distribution of stroke lesions in the ATLAS dataset. **Funding:** This work was supported by European Research Council Starting Grant 677697 (J.E.I.), Alzheimer's Research UK Interdisciplinary Grant ARUK-IRG2019A-003 (J.E.I.), National Institute on Aging grant R01AG070988 (J.E.I.), BRAIN Initiative grant RF1MH123195 (J.E.I. and B.F.), NIH Director's Office grant DP2HD101400 (B.L.E.), National Institute of Aging grant P30AG062421 (S.D., C.M., and S.E.A.), and National Institute of Biomedical Imaging and Bioengineering grant P41EB015902 (P.G.). **Author contributions:** Conceptualization: J.E.I., B.B., Y.B., B.L.E., D.C.A., P.G., and B.F. Methodology: J.E.I., B.B., Y.B., D.C.A., P.G., and B.F. Software: J.E.I., B.B., and Y.B. Investigation: J.E.I., B.B., C.M., S.E.A., and S.D. Resources: C.M., S.E.A., S.D., and B.F. Data curations: C.M., S.E.A., S.D., and B.F. Visualization: J.E.I. and B.B. Writing—original draft: J.E.I. and B.F. Writing—review and editing: B.B., Y.B., C.M., S.E.A., S.D., B.L.E., D.C.A., P.G., and B.F. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The MGH dataset (including MRI scans, ages, and genders) is available at the public server of the Martinos Center for Biomedical Imaging at https://ftp.nmr.mgh.harvard.edu/pub/dist/cnpublic/dist/MGH_dataset_wild_Iglesias_2022/. We note that the extracerebral tissue has been masked out in the images to prevent facial reconstruction and preserve anonymity. The rest of datasets (ATLAS, ADNI, BraTS, and BraTS-Reg) are publicly available. The source code of SynthSR is available on <https://github.com/BBillot/SynthSR>; we have included a current snapshot of the repository (for reproducibility purposes) at the same

address as above. A ready-to-use implementation of SynthSR (including the trained neural network) is distributed with FreeSurfer; please see <https://surfer.nmr.mgh.harvard.edu/fswiki/SynthSR>. We note that the training data cannot be released because it was collected under a set of Institutional Review Boards that did not include a provision approving public distribution.

Submitted 7 June 2022
Accepted 4 January 2023
Published 1 February 2023
10.1126/sciadv.add3607