



OPEN

Deep learning-Based 3D inpainting of brain MR images

Seung Kwan Kang^{1,7}, Seong A. Shin^{1,7}, Seongho Seo², Min Soo Byun³, Dong Young Lee⁴, Yu Kyeong Kim⁵, Dong Soo Lee⁶ & Jae Sung Lee^{1,6}✉

The detailed anatomical information of the brain provided by 3D magnetic resonance imaging (MRI) enables various neuroscience research. However, due to the long scan time for 3D MR images, 2D images are mainly obtained in clinical environments. The purpose of this study is to generate 3D images from a sparsely sampled 2D images using an inpainting deep neural network that has a U-net-like structure and DenseNet sub-blocks. To train the network, not only fidelity loss but also perceptual loss based on the VGG network were considered. Various methods were used to assess the overall similarity between the inpainted and original 3D data. In addition, morphological analyzes were performed to investigate whether the inpainted data produced local features similar to the original 3D data. The diagnostic ability using the inpainted data was also evaluated by investigating the pattern of morphological changes in disease groups. Brain anatomy details were efficiently recovered by the proposed neural network. In voxel-based analysis to assess gray matter volume and cortical thickness, differences between the inpainted data and the original 3D data were observed only in small clusters. The proposed method will be useful for utilizing advanced neuroimaging techniques with 2D MRI data.

The detailed anatomical information of the brain provided by T1-weighted magnetic resonance (MR) imaging with high-resolution three-dimensional (3D) sequences enables a variety of brain imaging studies, including quantitative measurements of brain tissue volume and cortical thickness, and classification of images for early diagnosis of disease^{1–6}. Moreover, 3D MR images are useful for accurate anatomical localization and quantification of co-registered functional images such as those obtained using positron emission tomography or single-photon emission tomography^{7,8}. However, 3D MRI sequences require longer scan times and larger data storage, so 2D images (multiple 2D slices with substantial gaps between them) are commonly acquired for diagnostic purposes in routine clinical practice. This may hinder the use of advanced neuroimage analysis techniques such as voxel-based morphometry and cortical thickness measurement for clinical data evaluation.

Numerous deep learning-based approaches have been proposed in recent years to solve a variety of computer vision problems, with significant results outperforming conventional algorithms^{9–14}. This is mainly because the improvement in computational power makes it possible to deal with large data sets based on highly manipulated neural network structures, thereby allowing them to solve complex problems. One of the advantages of deep learning over traditional mathematical and statistical methods is that it is relatively simple to configure the relationship between the input and the output spaces in various situations. Therefore, these methods have been quickly implemented in medical imaging and have achieved remarkable achievements in image classification, segmentation, noise reduction, and image generation^{15–26}.

In this study, we propose a deep learning-based inpainting technique that can generate 3D T1-weighted MR images from sparsely acquired 2D MR images (Fig. 1). The proposed network was trained to produce 3D inpainted MRI from the input that is linearly interpolated from sparsely sampled MR images. This approach allows the generation of 3D images without careful modeling of hand-crafted prior or assumptions about brain anatomy. The similarity between the inpainted and reference images was quantitatively analyzed based on real 3D MR images. We also used a voxel-based morphometry analysis and cortical thickness measurement to assess whether the complex structures of the cerebral cortex were correctly recovered and whether the inpainted data yielded equivalent morphological measurements to the original 3D data (reference image). Furthermore, we

¹Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Korea. ²Department of Electronic Engineering, Pai Chai University, Daejeon, Korea. ³Institute of Human Behavioral Medicine, Medical Research Center, Seoul National University, Seoul, Korea. ⁴Department of Psychiatry, Seoul National University College of Medicine, Seoul, Korea. ⁵Department of Nuclear Medicine, SMG-SNU Boramae Medical Center, Seoul, Korea. ⁶Department of Nuclear Medicine, Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea. ⁷These authors contributed equally: Seung Kwan Kang and Seong A. Shin. ✉email: jaes@snu.ac.kr

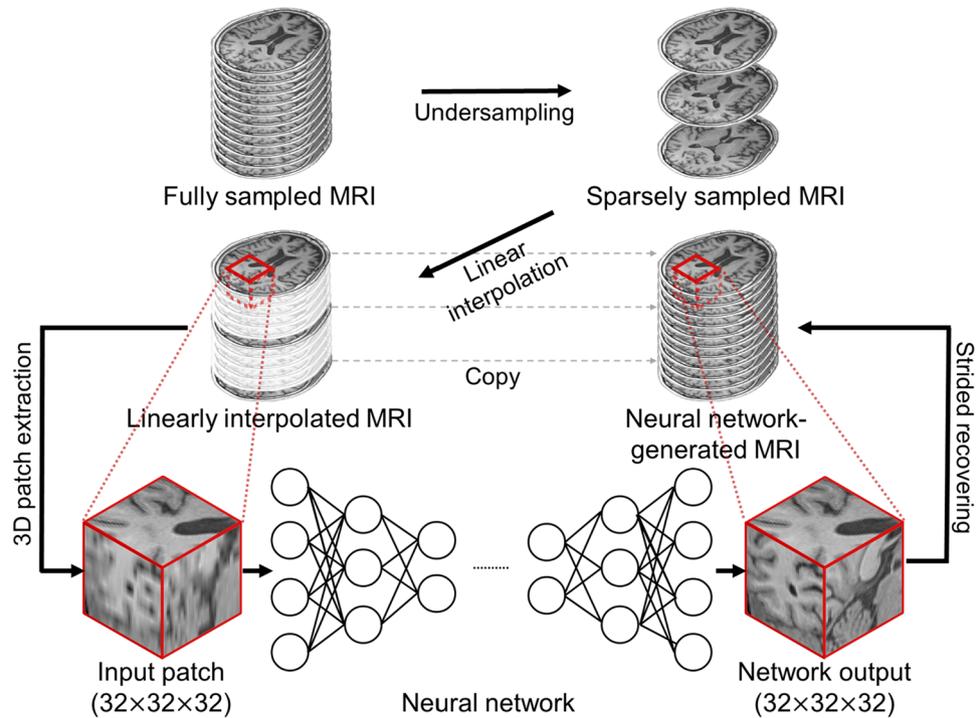


Figure 1. Schematic diagram of the proposed deep learning-based 3D inpainting method for magnetic resonance images.

Training set	NL	MCI	AD
<i>N</i>	338	117	72
Age (years)	62.4 ± 15.3	73.68 ± 6.9	72.56 ± 8.2
Gender (% female)	51.2%	65.0%	69.4%
Test set			
<i>N</i>	97	36	20
Age (years)	71.6 ± 8.0	72.7 ± 6.3	73.55 ± 8.6
Gender (% female)	60.8%	61.1%	65.0%

Table 1. Demographic characteristics of subjects: normal (NL) elderly, mild cognitive impairment (MCI) patients, and Alzheimer’s disease (AD) patients.

analyzed brain atrophy patterns in disease groups to see if we could characterize disease patterns using images generated based on deep learning.

Methods

Datasets. We trained and tested deep neural networks using 680 T1-weighted 3D MR images obtained from the Korean Brain Aging Study for Early Diagnosis and Prediction of Alzheimer’s Disease (KBASE)²⁷. Institutional Review Board of Seoul National University Hospital approved the study and all study participants provided informed consent. All experiments were conducted in accordance with relevant guidelines and regulations. T1-weighted 3D MR images of the subjects were obtained using a PET/MRI scanner (3-T, Biograph mMR, Siemens, Knoxville, Tennessee, USA). The imaging parameters were as follows: 208 slices with sagittal acquisition; slice thickness: 1 mm; repetition time: 1670 ms; echo time: 1.89 ms; flip angle: 9°; acquisition matrix: 208 × 256 × 256; and voxel size: 1.0 mm × 0.98 mm × 0.98 mm. Of the 680 datasets, 527 were used to train the neural network and the remaining 153 were used to test the trained network. Of the 153 datasets, 97 were normal (NL) elderly, 36 mild cognitive impairment (MCI) patients, and 20 Alzheimer’s disease (AD) patients (Table 1).

Neural network. *Proposed CNN architecture.* Figure 2 shows the detailed structure of the proposed network consisting of nine dense blocks²⁸ and transition layers (Supplementary Table 1). Each dense block with five convolutional layers is followed by a transition layer, yielding four blocks with a strided convolution transition layer, one without any transition layer, and four with a deconvolution transition layer. The convolutional layers

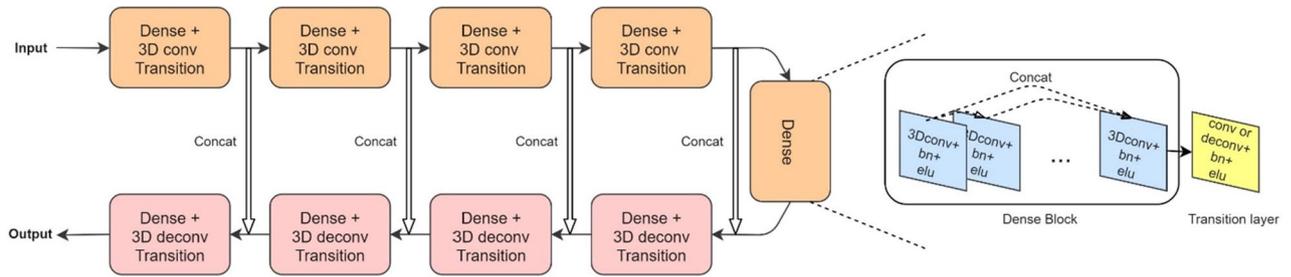


Figure 2. The architecture of the proposed neural network. A detailed explanation of the dense blocks and transition layers is shown on the right. All convolutions (conv) are calculated in a 3D manner, followed by batch renormalization (bn) and exponential linear units (elu) activation functions. The size of the input and output image patches is $32 \times 32 \times 32$. deconv: deconvolution. concat: concatenation.

retain the features of the previous layers by concatenating all the passing layers as described in the following equation:

$$\mathbf{x}_l = H_l([\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{l-1}]), \tag{1}$$

where \mathbf{x}_i represents the feature map of the i th layer and H_l represents the composition of the activation function (exponential linear unit) and the batch normalization of the l th layer²⁹. As shown in Fig. 2 and Eq. (1), each convolution operation in a dense block contains all previous outputs through concatenation and produces the same number of channels, so concatenation from all previous layers increases the number of input channels in the next layer. Therefore, we applied a $1 \times 1 \times 1$ convolution followed by a $3 \times 3 \times 3$ convolution to compress the data. The first dense block layer has 16 channels. The transition layers after the first four dense blocks subsample the images using strided ($2 \times 2 \times 2$) convolution to achieve a larger receptive field, whereas in the last four dense blocks, deconvolution (transposed convolution) transition layers follow the convolutional blocks. Features computed from the strided convolutional stages of the dense blocks were concatenated with those from the deconvolution dense blocks to ensure that the same dimension is maintained in the convolutional stage, enabling better backpropagation of the gradients. The architecture described is similar to that of the well-known U-Net³⁰.

Training details. As shown in Fig. 1, the reference images were sampled every fifth slice in the axial direction and then linearly interpolated. For the training datasets, non-brain tissue was excluded using a brain extraction tool (<http://www.nitrc.org/projects/mricron>). Input voxel patches of size $32 \times 32 \times 32$ were extracted from the linearly interpolated data for training with a stride of 16 if the center of the patch was in the extracted brain region. We normalized each MR image intensity to a range of -1 to 1 ^{31,32}. The normalization factor was saved and used to rescale the image after reconstructing 3D image. A total of 439,479 training patches were used for network training with a batch size of 12. After training, the neural network was evaluated using test patches extracted with a stride of 24 and eight overlapping voxels. The overlapping voxels were averaged while reconstructing the inpainted 3D image from the patches. The initial learning rate was set to 0.0003, which decreased to 10% of the initial learning rate after 20 epochs. The network was trained for 60 epochs on a computer with an Intel i7-8770 processor and a Nvidia Geforce GTX 1080ti graphics card.

Loss functions. To train the network, not only fidelity loss but also perceptual loss based on VGG network were considered³¹. Accordingly, the total loss was:

$$L = L_{fid} + \gamma_1 L_{per} \tag{2}$$

where,

$$L_{fid} = \frac{1}{mn_v} \sum_{i=1}^m \|\mathbf{y}^i - f(\mathbf{x}_s^i)\|, \tag{3}$$

$$L_{per} = \frac{1}{m} \sum_{i=1}^m \left\| \phi(\mathbf{y}^i) - \phi(f(\mathbf{x}_s^i)) \right\|_2^2, \tag{4}$$

m is the batch size, \mathbf{y} is the vector of reference data, \mathbf{x}_s is the sparsely sampled input, and f is the neural network. L_{fid} is the fidelity loss reducing the content difference between the reference and the neural network output, L_{per} is the perceptual loss in feature space³³, ϕ is the intermediate feature map of a particular neural network, and γ_1 is the tuning parameter for the loss functions. In this study, the fourth convolutional layer in the VGG19 network was used as ϕ feature map¹². Given that the VGG19 network requires a single RGB image composed of three channels as input, the 16th sagittal, coronal, and transaxial slices of the 3D input patch were supplied as inputs to the pre-trained VGG19 network. The perceptual loss was evaluated using only the central slices of the input patch for fast network training. The input to the VGG19 was upsampled from 32×32 pixels to 224×224

pixels using a linear interpolation method so that the size of the extracted slices matched that of VGG19. The tuning parameter γ_1 was set to 10^{-6} .

Comparison. For comparison, we also implemented the well-known U-Net using the same numbers of convolutional filters and pooling layers as described in the original paper³⁰, but using a different convolution operation dimension (3D instead of 2D) of and input and output sizes (both $32 \times 32 \times 32$). The U-Net was trained and tested using the same dataset used for the proposed network.

Similarity evaluations. For the test dataset, three quantitative measures were used to evaluate the similarity between the reference image (original 3D image) and the inpainted images (linear interpolated data and neural network-generated data). The similarity assessment was applied to the brain regions defined by the brain mask produced using the brain extraction tool³⁴. Peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and high-frequency error norm (HFEN) were calculated using the following equations^{35,36}.

$$PSNR = 10 \log_{10} \left(\frac{\max(I_y)}{MSE} \right), \quad (5)$$

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (6)$$

$$HFEN = \frac{\|LoG(I_x) - LoG(I_y)\|_2}{\|LoG(I_y)\|_2}, \quad (7)$$

where I_x is the test image, I_y is the reference image, and MSE is the mean squared error between them. Here, $\mu_{(\bullet)}$ and $\sigma_{(\bullet)}$ denote the mean and variance or covariance of two images, and $c_1 = (0.01 \times d)^2$ and $c_2 = (0.03 \times d)^2$. $LoG(\bullet)$ denotes the Laplacian of 3D Gaussian filter function with the kernel size of $15 \times 15 \times 15$.

Voxel-based morphometry analysis and cortical thickness measurements. Different brain tissue types (gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF)) were segmented from the reference and inpainted images using Statistical Parametric Mapping 12 (SPM12; <http://www.fil.ion.ac.uk/spm>)³⁷. The segmented images were then quantitatively examined using Dice similarity coefficient³⁸ to compare the similarity between the reference and inpainted data.

To assess the accuracy of the volumetric measures estimated from the inpainted data, a voxel-based morphometry method based on the SPM12 pipeline was applied to the GM images. Inter-subject registration of GM images was performed based on nonlinear deformation using a fast diffeomorphic image registration toolbox³⁹, and the images were modulated to preserve the tissue volume. The images were then smoothed with a Gaussian kernel with a full width at half maximum of 10 mm. The GM volume images of the inpainted data were then compared to the reference data in a voxel-wise manner using t -test with the total intracranial volume as a covariate. The statistical threshold was set to p -value of 0.005 (corrected for family-wise error) and cluster extent of 100 voxels. The GM volumes in preselected regions (frontal, lateral parietal, posterior cingulate-precuneus, occipital, lateral temporal, hippocampus, caudate, and putamen regions) were also estimated using an automatic anatomic labelling algorithm^{40,41} to evaluate the correlation between the inpainted and reference data.

Cortical thickness measurements were performed using the automated image preprocessing pipeline of FreeSurfer (v6.0.0; <http://surfer.nmr.mgh.harvard.edu/>). The processing pipeline included skull stripping, nonlinear registration, subcortical segmentation, cortical surface reconstruction, and parcellation⁴²⁻⁴⁵. The output of the automated processes was visually inspected for obvious errors. Instead of manually correcting for the errors, 15 subjects were excluded from the statistical analysis. Cortical thickness values were smoothed via surface-based smoothing with a Gaussian kernel with a full width at half maximum of 5 mm. Moreover, a two-sample t -test was performed to detect brain regions with differences in cortical thickness between the inpainted and reference data. Results were considered significant if survived at false-discovery rate corrected p -value of 0.05 at cluster level. The correlation of cortical thickness between the inpainted and reference data was also analyzed.

Finally, we performed voxel-wise comparisons of the GM volume and cortical thickness between NL and MCI/AD groups to investigate whether the inpainted images yield similar patterns of morphological abnormalities similar to those shown by the reference data in the MCI and AD groups.

Results

Similarity between inpainted and reference images. The left columns of Fig. 3a show the transaxial slices of the neural network-generated data (third and fourth rows) and the linearly interpolated data⁴⁶ (second row) compared to the reference data (top row). The proposed neural network has well recovered the anatomical details in the brain cortices and ventricles and showed high contrast between GM and WM. In particular, the output of the proposed method had less blurred structures and edges than that of U-Net. On the other hand, the linear interpolation method led to interpolation artifacts and image blurring. The artifacts and blurring most pronounced around the ventricles have resulted in high distortion in the coronal and sagittal planes.

PSNR, SSIM, and HFEN between the inpainted and reference data were calculated in a 3D manner for 153 test datasets. As shown in Fig. 3b-d, all metrics indicate that the images generated using the proposed neural

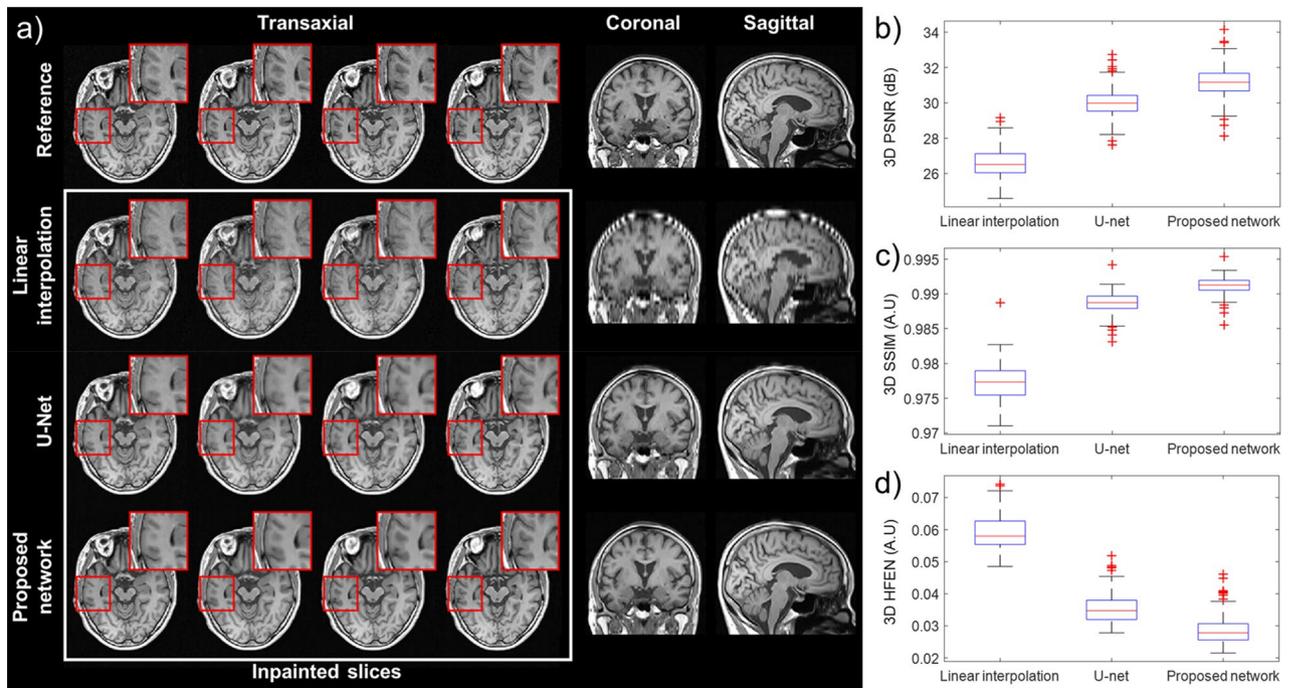


Figure 3. Comparison of inpainted images to the original 3D MR images (reference). (a) Each column shows the 137th to 140th transaxial slices, the 95th coronal slice, and the 125th sagittal slice. Results of PSNR (b), SSIM (c) and HFEN (d) evaluation of the accuracy of the different inpainting methods (linear interpolation, U-Net and proposed network) relative to reference data are shown on the left.

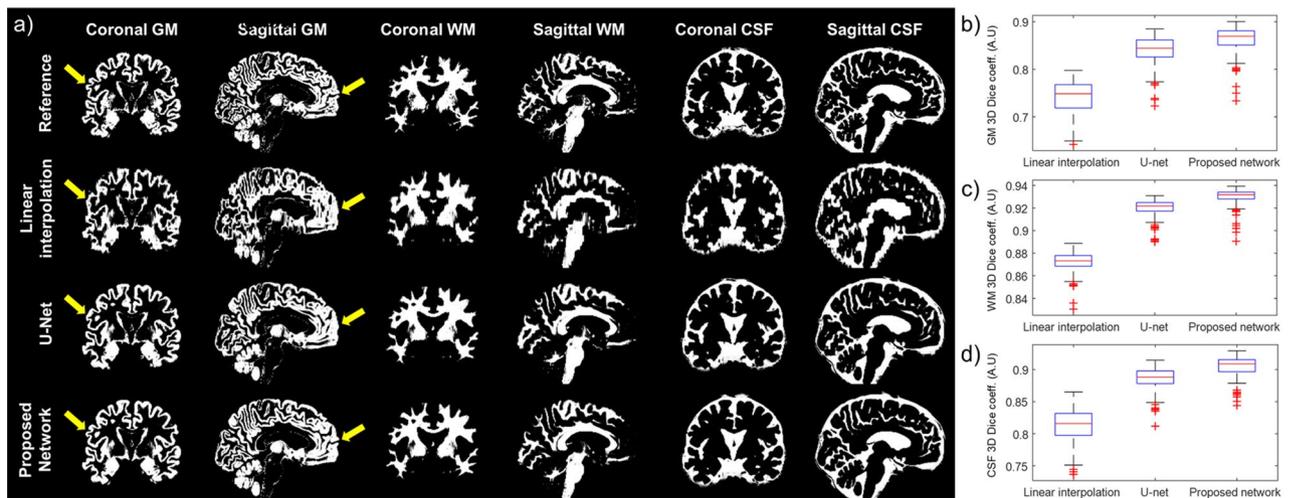


Figure 4. Gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) segments of a representative image. (a) The areas indicated by arrows show poor recovery of the GM segments in linearly interpolated images (second row) and U-Net images (third row) compared to the proposed neural network-generated images (fourth row). 3D Dice coefficients between reference (original 3D MRI) and inpainted data in GM (b), WM (c), and CSF (d) are shown on the left.

network method are more similar to the reference than the images generated using the conventional linear interpolation method and U-Net.

The superiority of the proposed method was also evident in the segmentation of GM in the inpainted images. Figure 4a shows the coronal and sagittal planes of segmented GM, WM, and CSF images. In the GM maps, the cortical gyri were not fully recovered when using linear interpolation and U-Net as indicated by the arrows in the figure. However, the GM maps generated by the proposed neural network method are almost identical to the reference images. Figure 4b–d show the Dice similarity coefficients relating to the inpainted and reference images for the different tissue maps, indicating that the proposed method outperformed the conventional interpolation method and U-Net.

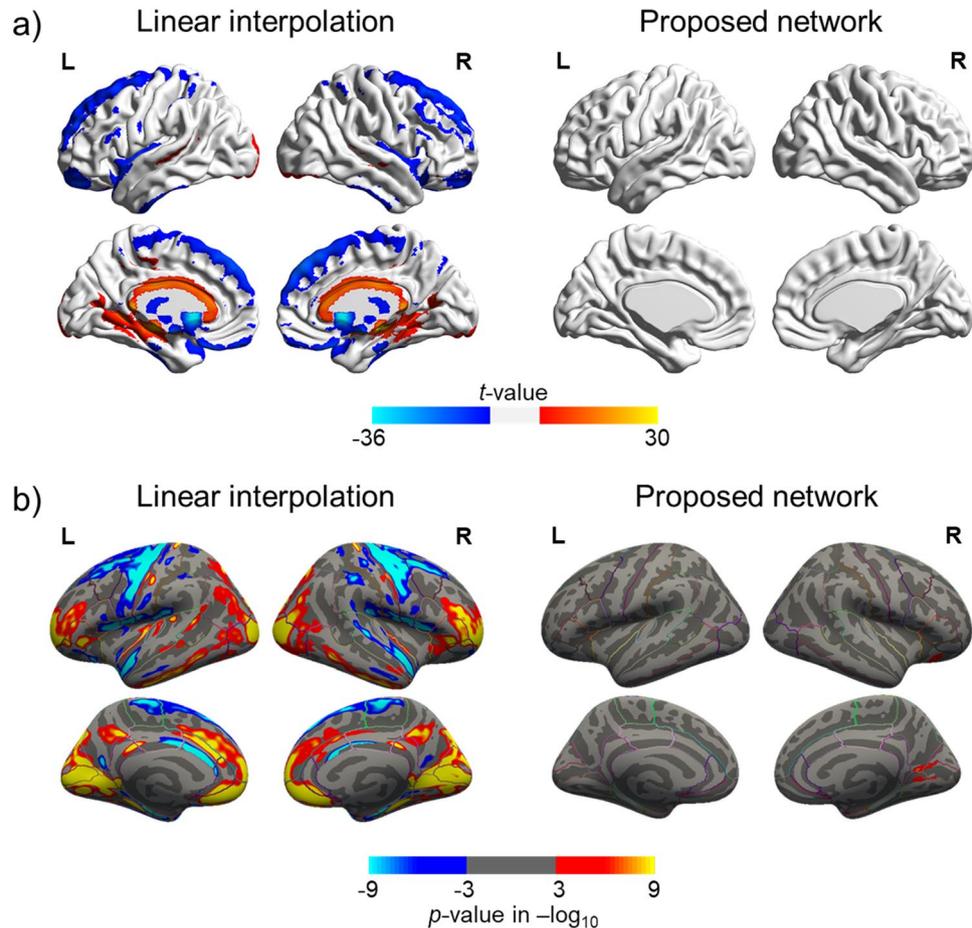


Figure 5. Voxel-wise comparison of inpainted data to the reference data (original 3D MRI). **(a)** Gray matter volume difference. **(b)** Cortical thickness difference. Red indicates overestimation, and blue indicates underestimation relative to the reference data.

Voxel-based morphometry and cortical thickness analyses. Compared to the reference data, the proposed neural network-generated inpainted data showed no significant differences in the GM volume estimates (Fig. 5a and supplementary Fig. 1). On the other hand, the linearly interpolated data showed overestimation in the medial areas of the brain and underestimation in the bilateral frontal and occipital lobe, insula, striatum, and thalamus with respect to the GM volume. Similarly, as shown in Fig. 5b, when using the neural network-generated data, differences in the cortical thickness were found only in the small clusters in the medial occipital lobe. On the other hand, the cortical thickness was overestimated in the cingulate, bilateral occipital lobe, prefrontal cortex, and temporal lobe and underestimated in the bilateral frontal lobe and lateral temporal lobe when using the linearly interpolated data.

The individual regional GM volumes estimated using the neural network-generated data are highly correlated with the volumes estimated using the reference data (Fig. 6a). For all the regions analyzed in the study, the volumetric measures obtained using the neural network-generated data showed a higher correlation than the measures obtained with the linearly interpolated data. Regional cortical thicknesses obtained with the inpainted and reference data were also compared (Fig. 6b). The correlation between the neural network-generated and reference data was higher in all regions.

Figure 7a and Supplementary Fig. 2 shows the patterns of the pathological gray matter volume reductions in the MCI and AD groups compared to the NL group identified using the inpainted and reference data. The reference data and inpainted data based on the proposed network showed GM volume reductions in the bilateral medial temporal lobe in the MCI group, whereas the bilateral frontal, parietal, and temporal lobes were significantly affected in the AD group. As shown in Fig. 7b, the atrophy patterns observed in the MCI and AD groups using the neural network-generated data were more similar to the patterns found using the reference data. In both the inpainted data using the proposed method and reference data, the cortical thickness of the AD group was significantly decreased in the temporal, parietal, dorsolateral frontal, and posterior cingulate cortices. However, the thinning of the cortex in the MCI group was almost undetectable.

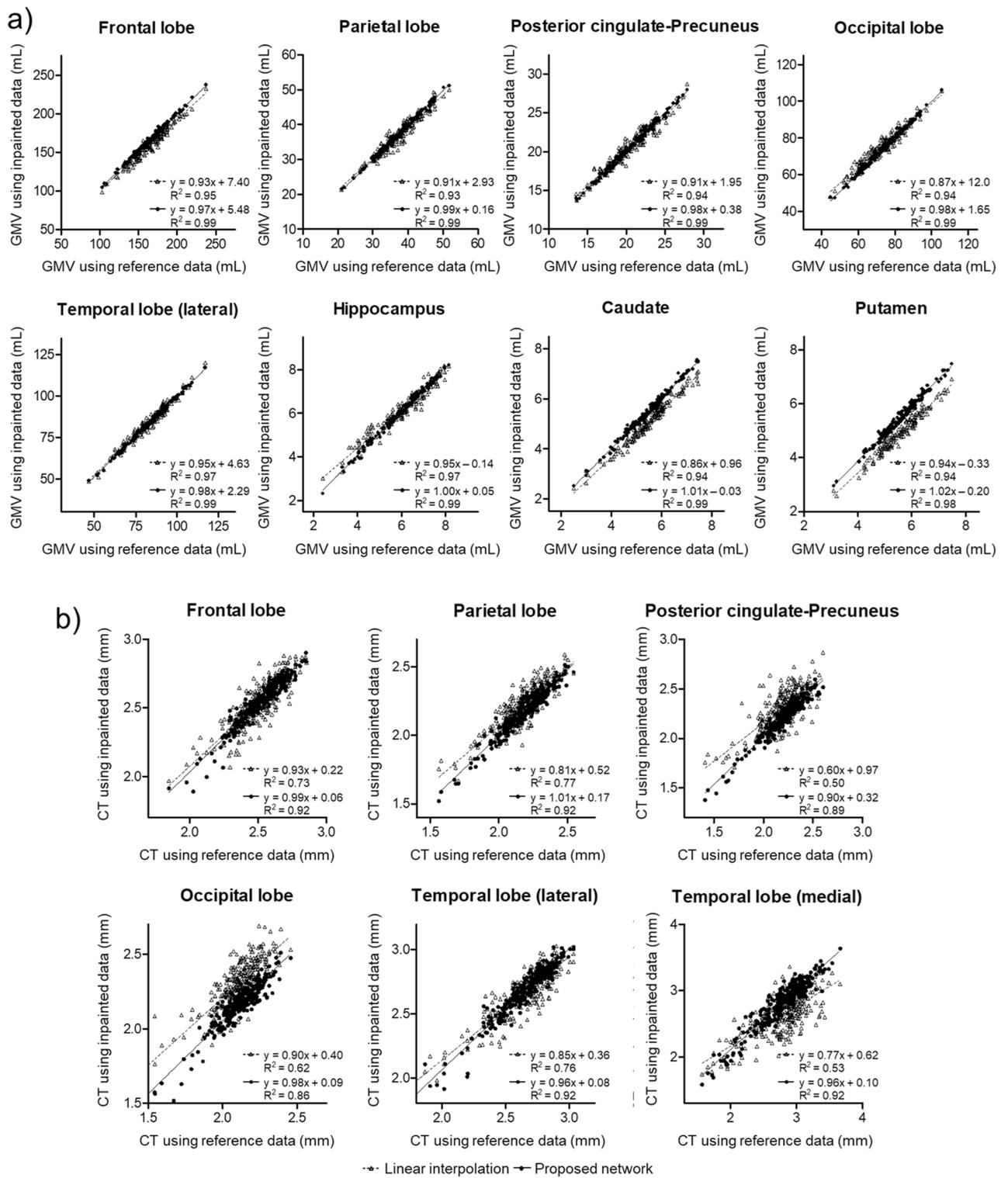


Figure 6. Correlations of regional gray matter volume (a) or cortical thickness (b) between reference (x-axis) and inpainted (y-axis) data. Solid lines with a filled circle are for the proposed neural network, and dashed lines with an empty triangle are for linear interpolation.

Discussion

This study proposes a deep neural network approach for the 3D inpainting of sparsely sampled human brain MR images. We observed that the brain structure was well recovered using the proposed neural network method without significant interpolation artifacts or noise. Several similarity measures showed that the inpainted images obtained using the proposed method were more similar to the original reference images than those obtained using the linear interpolation method and U-Net. Furthermore, voxel-wise comparison analyses of the GM

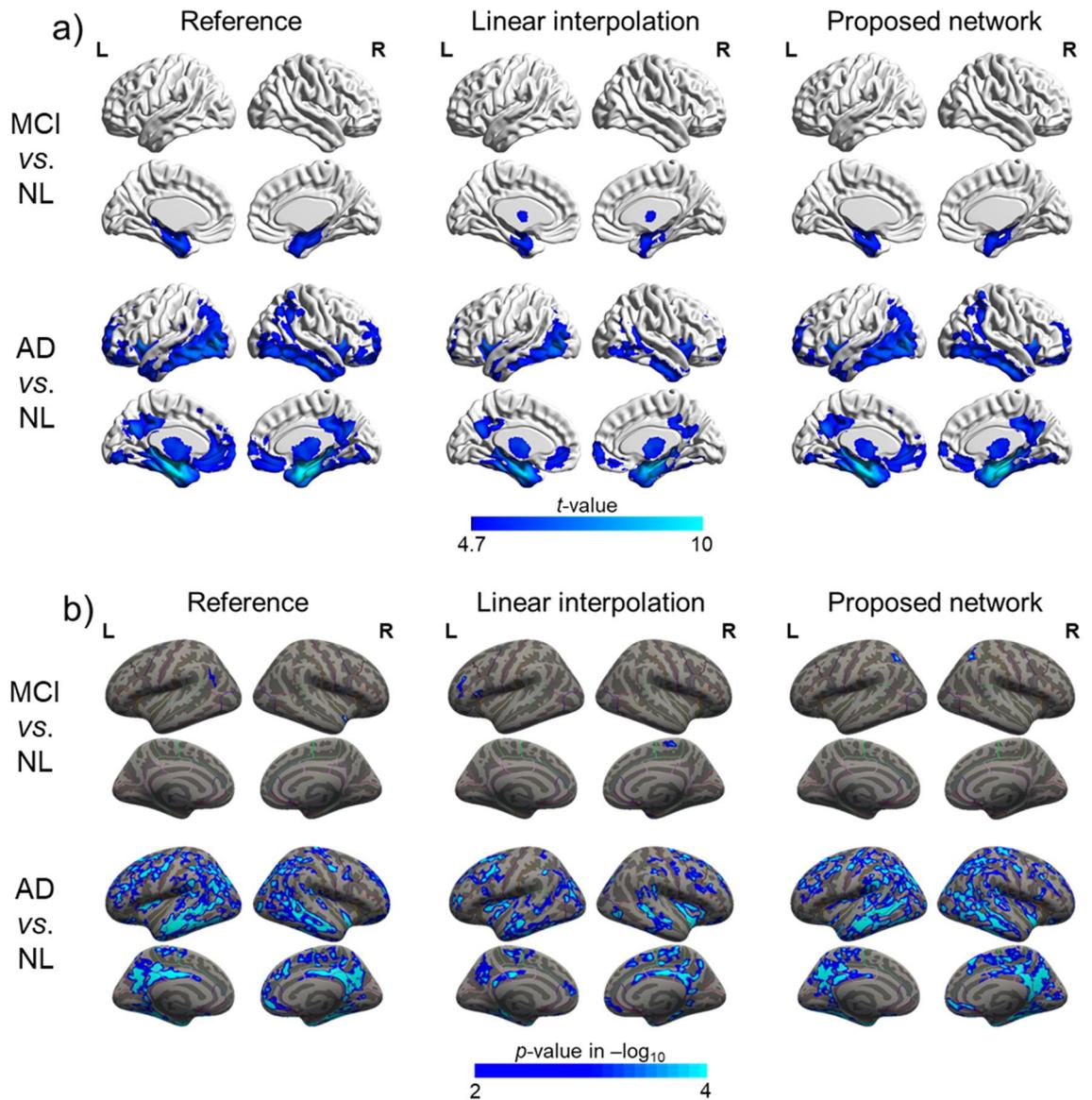


Figure 7. Group analysis among NL, MCI and AD. **(a)** Gray matter volume. **(b)** Cortical thickness.

volume and cortical thickness demonstrated that the 3D MR images generated using the proposed neural network outperformed the conventional linearly interpolated images in morphometric measure estimation. Additionally, the cortical atrophy patterns in patient groups identified using neural network-generated data resembled the patterns found using the reference data, reflecting the potential implications of the proposed method for clinical and neuroimaging research.

In comparison with previously suggested medical image interpolation and super resolution technologies^{47–51}, the proposed method has the advantage that no prior knowledge regarding the body or organ shapes, or other additional information is required. The methods based on partial differential equations, such as the one proposed by Li et al. assume that the objects reconstructed from the sparsely sampled slices are smooth⁵¹. In addition, these methods typically require an iterative updating process and careful tuning of the regularization parameters. Integrating information obtained from multiple pulse sequences is another widely used method for conducting inpainting or performing super-resolution in MR images. The subpixel shift with multiple acquisitions and iterative back projection is an example of this approach⁴⁷. Another proposed method employs non-local means between prior high-resolution images and nearest-neighbor interpolated low-resolution images acquired using multiple different pulse sequences⁴⁸. A recently proposed method employs a sparse representation framework using multiple orthogonal anisotropic resolution scans^{49,50}. On the contrary, the proposed deep learning approach yields 3D MR images only from the sparsely sampled 2D images given as an input to the trained neural network. Although the proposed network conducted successful inpainting and yielded better performance than U-Net, the current setup might be sub-optimal because numerous network models are surpassing the U-Net in many other applications^{28,52–54}. Comprehensive comparisons with state-of-the-art networks are needed as future work.

The voxel-based analyses of the GM volume and cortical thickness offers the reliability of employing inpainted images for neuroimaging research. Moreover, using the network-generated images, the individual regional GM volume estimates including the subcortical structures highly correlated with the GM volume estimates obtained using the reference data. Subcortical structures, such as the hippocampus and striatum, generally show high inter-individual anatomical variability and are highly affected in diseased conditions, which makes them prone to produce reconstruction errors, diminishing accuracy in 3D interpolation processes. Our results demonstrated that the recovery of the 3D brain images using the proposed method is reliable for clinical and neuroimaging research applications. The acquisition of 3D MR images using the proposed method offers several advantages when only 2D slices are available. The utilization of 3D MR images facilitates the diagnosis of neurodegenerative diseases since regional volume measurements often serve as an important neuroimaging biomarker in many neurodegenerative diseases (for example, hippocampal volume in dementia and striatal volume in movement disorders^{55–58}). Moreover, 3D images can aid in monitoring disease progression and predicting prognosis. The pattern of progressive cerebral volume reduction, which precedes the onset of a disease, can be detected using the inpainted 3D images^{56,59}. For example, cerebral atrophy pattern lies in line with the pathological staging scheme in Alzheimer's disease, and allows the early detection of dementia and prediction of the disease progression. The inpainting of brain images can also be used for 3D MRI-guided neurosurgical planning and intraoperative navigation as they enable 3D visualization of a surgical area, whereas 2D slices only offer a limited view. Using 3D data, morphometric measurements, such as cortical thickness and tissue volume, can be performed and assessed as shown in this study. Furthermore, the 3D inpainting technique would enable the application of recently proposed automated methods used to classify brain images to ascertain the type of disease as well as for the prediction of diseases based on machine learning.

Although the proposed 3D SR method could recover brain structures in unscanned slices well, further investigation regarding its ability to detect small lesions should be conducted. However, it is unlikely that the proposed method can recover lesions that are not included in the scanned 2D slices. Supplementary Fig. 3 shows an example of how we can utilize the proposed method in the clinical setting. In the typical simultaneous head-and-neck PET/MRI studies, only the sparsely sampled transaxial 2D MRI slices are acquired although PET images are reconstructed with almost isotropic voxel size. The 3D MR images with the same voxel size as PET generated using the proposed method will be useful for the anatomical localization of PET findings in MR images and other various tasks, such as tumor volume estimation and MRI-guided PET reconstruction.

As a future study, further investigation on the generalization power of the proposed method across sites and countries will be necessary. Supplementary Fig. 4 shows the inpainting results of T2* and FLAIR MR images available in the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. They were obtained by applying the proposed deep neural network trained using only the T1 MR images obtained from our institute, highlighting the highly potential generalization power of our method.

In this study, we fixed the sampling interval and offset to five slices and one slice, respectively. The effect of offset (the first slice number for slice sampling) doesn't seem to be considerable as shown in Supplementary Fig. 5, which depicts the PSNR values obtained from tests conducted with the networks trained with a single offset (= 1) and multiple other offsets (= 1, 2, ..., 5). Further studies based on transfer learning for different numbers of interleaving slices will make this method more useful. In this study, we only considered filling the gaps of transaxial slices, but applications in other directions are also possible. A simple way to deal with various missing data conditions is to train the network with a mixed data set containing different intervals in different directions.

A limitation of the proposed method is the boundary artifacts caused by patch-based learning. Our task was to fill the missing transaxial slices, and unwanted boundaries were observed in the coronal and sagittal planes. The proposed network showed fewer artifacts than the U-net (Supplementary Fig. 6). A potential reason is that the given slices copied from the reference image did not produce a difference. Moreover, the difference between training and test set can make the problem worse. To reduce this artifact, we can use more patches with a smaller number of strides. Alternatively, we can have a larger number of voxels overlap to make the boundary smoother. However, the former option can increase training inefficiencies, resulting in extended learning time, and the latter requires more time to reconstruct the 3D images from 2D slices. There is a recent report that the generative model and attention gates applied to the inpainting of natural images reduced boundary artifacts compared to the deterministic model⁵³. This strategy could also be used in medical imaging to reduce these boundary artifacts.

The number of the training set was selected heuristically (527 out of 680 images). This is another limitation of this study that can be improved with rigorous dataset determination. For example, we can examine the age and diagnosis of the entire dataset and divide it into training and test set to have similar distributions. We only used three central slices of the input patch to calculate the perceptual loss because the VGG network used for assessing the perceptual loss receives only three-channel 2D images as input. Although the incorporation of perceptual loss sharpened the edges of the network output, the effect was not dominant. We need to investigate whether the perceptual loss based on a network trained in 3D will improve network performance.

Conclusions

In this study, we proposed a deep neural network for generating inpainted 3D MR images from intermittently sampled 2D images. Patch-based training of the network was performed by simultaneously minimizing the fidelity and perceptual losses for a large dataset. Through quantitative and qualitative evaluation, it was shown that the proposed method outperformed the conventional linear interpolation method in terms of the similarity between the inpainted and reference images and the validity of the inpainted images in the morphometric measurements.

Data availability

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Received: 23 December 2019; Accepted: 14 December 2020

Published online: 18 January 2021

References

- Korolev, I. O., Symonds, L. L. & Bozoki, A. C. Predicting progression from mild cognitive impairment to Alzheimer's dementia using clinical, MRI, and plasma biomarkers via probabilistic pattern classification. *PLoS ONE* **11**, e0138866. <https://doi.org/10.1371/journal.pone.0138866> (2016).
- Kloppel, S. *et al.* Automatic classification of MR scans in Alzheimer's disease. *Brain* **131**, 681–689. <https://doi.org/10.1093/brain/awm319> (2008).
- Koikkalainen, J. *et al.* Differential diagnosis of neurodegenerative diseases using structural MRI data. *NeuroImage. Clinical* **11**, 435–449. <https://doi.org/10.1016/j.nicl.2016.02.019> (2016).
- Koutsouleris, N. *et al.* Individualized differential diagnosis of schizophrenia and mood disorders using neuroanatomical biomarkers. *Brain* **138**, 2059–2073. <https://doi.org/10.1093/brain/awv111> (2015).
- Bendfeldt, K. *et al.* Multivariate pattern classification of gray matter pathology in multiple sclerosis. *Neuroimage* **60**, 400–408. <https://doi.org/10.1016/j.neuroimage.2011.12.070> (2012).
- Marzelli, M. J., Hoeff, F., Hong, D. S. & Reiss, A. L. Neuroanatomical spatial patterns in turner syndrome. *Neuroimage* **55**, 439–447. <https://doi.org/10.1016/j.neuroimage.2010.12.054> (2011).
- Friston, K. J. *et al.* Spatial registration and normalization of images. *Hum. Brain Mapp.* **3**, 165–189. <https://doi.org/10.1002/hbm.460030303> (1995).
- Maes, F., Collignon, A., Vandermeulen, D., Marchal, G. & Suetens, P. Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imaging* **16**, 187–198. <https://doi.org/10.1109/42.563664> (1997).
- He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. [arXiv:1512.03385](https://arxiv.org/abs/1512.03385) (2015).
- Krizhevsky, A., Sutskever, I. & Hinton, G. E. (2012) Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* 1097–1105
- Long, J., Shelhamer, E. & Darrell, T. Fully convolutional networks for semantic segmentation. In *CVPR*, 3431–3440 (2015).
- Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2014).
- Xie, J., Xu, L. & Chen, E. Image denoising and inpainting with deep neural networks. *Adv. Neural Inf. Process. Syst.* **25**, 341–349 (2012).
- Lee, J. S. A review of deep learning-based approaches for attenuation correction in positron emission tomography. *IEEE Trans. Radiat. Plasma Med. Sci.* <https://doi.org/10.1109/TRPMS.2020.3009269> (2020).
- Brebisson, A. D. & Montana, G. Deep neural networks for anatomical brain segmentation. [arXiv:1502.02445](https://arxiv.org/abs/1502.02445) (2015).
- Ciresan, D., Giusti, A., Gambardella, L. M. & Schmidhuber, J. Deep neural networks segment neuronal membranes in electron microscopy images. *Adv. Neural Inf. Process. Syst.* **25**, 2843–2851 (2012).
- Dey, D., Chaudhuri, S. & Munshi, S. Obstructive sleep apnoea detection using convolutional neural network based deep learning framework. *Biomed. Eng. Lett.* **8**, 95–100. <https://doi.org/10.1007/s13534-017-0055-y> (2018).
- Hwang, D. *et al.* Improving the accuracy of simultaneously reconstructed activity and attenuation maps using deep learning. *J. Nucl. Med.* **59**, 1624–1629. <https://doi.org/10.2967/jnumed.117.202317> (2018).
- Kang, S. K. *et al.* Adaptive template generation for amyloid PET using a deep learning approach. *Hum. Brain Mapp.* **39**, 3769–3778. <https://doi.org/10.1002/hbm.24210> (2018).
- Leynes, A. P. *et al.* Direct PseudoCT generation for pelvis PET/MRI attenuation correction using deep convolutional neural networks with multi-parametric MRI: zero echo-time and Dixon deep pseudoCT (ZeDD-CT). *J. Nucl. Med.* **59**, 852–858. <https://doi.org/10.2967/jnumed.117.198051> (2017).
- Mansour, R. F. Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy. *Biomed. Eng. Lett.* **8**, 41–57. <https://doi.org/10.1007/s13534-017-0047-y> (2018).
- Nie, D., Cao, X., Gao, Y., Wang, L. & Shen, D. Estimating CT Image from MRI data using 3D FULLY CONVOLUTIONAL NETWORKS. In *MICCAI*, 170–178 (2016).
- Park, J. *et al.* Computed tomography super-resolution using deep convolutional neural network. *Phys. Med. Biol.* **63**, 145011. <https://doi.org/10.1088/1361-6560/aacdd4> (2018).
- Suk, H.-I., Lee, S.-W., Shen, D. & Initiative, A. S. D. N. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *Neuroimage* **101**, 569–582 (2014).
- Hwang, D. *et al.* Generation of PET attenuation map for whole-body time-of-flight 18F-FDG PET/MRI using a deep neural network trained with simultaneously reconstructed activity and attenuation maps. *J. Nucl. Med.*, jnumed. 118.219493 (2019).
- Park, J. *et al.* Measurement of glomerular filtration rate using quantitative SPECT/CT and deep-learning-based kidney segmentation. *Sci. Rep.* **9**, 4223 (2019).
- Byun, M. S. *et al.* Korean brain aging study for the early diagnosis and prediction of Alzheimer's disease: methodology and baseline sample characteristics. *Psychiatry Investig.* **14**, 851–863. <https://doi.org/10.4306/pi.2017.14.6.851> (2017).
- Huang, G., Liu, Z., Maaten, L. v. d. & Weinberger, K. Q. Densely connected convolutional networks. [arXiv:1608.06993](https://arxiv.org/abs/1608.06993) (2016).
- Clevert, D.-A., Unterthiner, T. & Hochreiter, S. Fast and accurate deep network learning by exponential linear units (ELUs). [arXiv:1511.07289](https://arxiv.org/abs/1511.07289) (2015).
- Ronneberger, O., Fischer, P. & Brox, T. 2015 U-net: convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp 234–241
- Radford, A., Metz, L. & Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. [arXiv:1511.06434](https://arxiv.org/abs/1511.06434) (2015).
- Ledig, C. *et al.* Photo-realistic single image super-resolution using a generative adversarial network. In *CVPR*, 4681–4690 (2017)
- Johnson, J., Alahi, A. & Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. In: *ECCV*, 694–711 (2016).
- Smith, S. M. Fast robust automated brain extraction. *Hum. Brain Mapp.* **17**, 143–155 (2002).
- Hore, A. & Ziou, D. Image quality metrics: PSNR vs. SSIM. In *ICPR*, 2366–2369 (2010).
- Ravishanker, S. & Bresler, Y. MR image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Trans. Med. Imaging* **30**, 1028–1041. <https://doi.org/10.1109/TMI.2010.2090538> (2011).
- Ashburner, J. & Friston, K. J. Unified segmentation. *Neuroimage* **26**, 839–851 (2005).
- Dice, L. R. Measures of the amount of ecologic association between species. *Ecology* **26**, 297–302 (1945).
- Ashburner, J. A fast diffeomorphic image registration algorithm. *Neuroimage* **38**, 95–113. <https://doi.org/10.1016/j.neuroimage.2007.07.007> (2007).
- Reiman, E. M. *et al.* Fibrillar amyloid-beta burden in cognitively normal people at 3 levels of genetic risk for Alzheimer's disease. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 6820–6825. <https://doi.org/10.1073/pnas.0900345106> (2009).

41. Tzourio-Mazoyer, N. *et al.* Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15**, 273–289 (2002).
42. Desikan, R. S. *et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968–980. <https://doi.org/10.1016/j.neuroimage.2006.01.021> (2006).
43. Fischl, B. *et al.* Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* **33**, 341–355 (2002).
44. Segonne, F. *et al.* A hybrid approach to the skull stripping problem in MRI. *Neuroimage* **22**, 1060–1075. <https://doi.org/10.1016/j.neuroimage.2004.03.032> (2004).
45. Segonne, F., Grimson, E. & Fischl, B. A genetic algorithm for the topology correction of cortical surfaces. *Inf. Process. Med. Imaging* **19**, 393–405 (2005).
46. Goshtasby, A., Turner, D. A. & Ackerman, L. V. Matching of tomographic slices for interpolation. *IEEE Trans. Med. Imaging* **11**, 507–516. <https://doi.org/10.1109/42.192686> (1992).
47. Greenspan, H., Oz, G., Kiryati, N. & Peled, S. MRI inter-slice reconstruction using super-resolution. *Magn. Reson. Imaging* **20**, 437–446. [https://doi.org/10.1016/S0730-725X\(02\)00511-8](https://doi.org/10.1016/S0730-725X(02)00511-8) (2002).
48. Jafari-Khouzani, K. MRI upsampling using feature-based nonlocal means approach. *IEEE Trans. Med. Imaging* **33**, 1969–1985. <https://doi.org/10.1109/tmi.2014.2329271> (2014).
49. Jia, Y., He, Z., Gholipour, A. & Warfield, S. K. Single anisotropic 3-D MR Image upsampling via overcomplete dictionary trained from in-plane high resolution slices. *IEEE J. Biomed. Health Inform.* **20**, 1552–1561. <https://doi.org/10.1109/JBHI.2015.2470682> (2016).
50. Jia, Y., Gholipour, A., He, Z. & Warfield, S. K. A new sparse representation framework for reconstruction of an isotropic high spatial resolution MR volume from orthogonal anisotropic resolution scans. *IEEE Trans. Med. Imaging* **36**, 1182–1193. <https://doi.org/10.1109/TMI.2017.2656907> (2017).
51. Li, Y., Lee, H. G., Xia, B. & Kim, J. A compact fourth-order finite difference scheme for the three-dimensional Cahn–Hilliard equation. *Comput. Phys. Commun.* **200**, 108–116 (2016).
52. Liu, G. *et al.* Image inpainting for irregular holes using partial convolutions. In *ECCV*, 85–100 (2018)
53. Yu, J. *et al.* Generative image inpainting with contextual attention. In *CVPR*, 5505–5514 (2018)
54. Yu, J. *et al.* Free-form image inpainting with gated convolution. In *ICCV*, 4471–4480 (2019)
55. Aylward, E. H. Change in MRI striatal volumes as a biomarker in preclinical Huntington’s disease. *Brain Res. Bull.* **72**, 152–158. <https://doi.org/10.1016/j.brainresbull.2006.10.028> (2007).
56. Ruan, Q. *et al.* Potential neuroimaging biomarkers of pathologic brain changes in mild cognitive impairment and Alzheimer’s disease: a systematic review. *BMC Geriatr.* **16**, 104–104. <https://doi.org/10.1186/s12877-016-0281-7> (2016).
57. Saeed, U. *et al.* Imaging biomarkers in Parkinson’s disease and Parkinsonian syndromes: current and emerging concepts. *Transl. Neurodegener.* **6**, 8. <https://doi.org/10.1186/s40035-017-0076-6> (2017).
58. van Tol, M. J. *et al.* Regional brain volume in depression and anxiety disorders. *Arch. Gen. Psychiatry* **67**, 1002–1011. <https://doi.org/10.1001/archgenpsychiatry.2010.121> (2010).
59. Mathalon, D. H., Sullivan, E. V., Lim, K. O. & Pfefferbaum, A. Progressive brain volume changes and the clinical course of schizophrenia in men: a longitudinal magnetic resonance imaging study. *Arch. Gen. Psychiatry* **58**, 148–157. <https://doi.org/10.1001/archpsyc.58.2.148> (2001).

Acknowledgements

This work was supported by grants from the National Research Foundation of Korea (NRF) funded by the Korean Ministry of Science and ICT (grant no. NRF-2014M3C7034000, NRF-2016R1A2B3014645, and 2014M3C7A1046042). The funding source had no involvement in the study design, collection, analysis, or interpretation.

Author contributions

S.K.K. designed and performed computational codes and numerical experiments, collected and analyzed data and drafted the manuscript; S.A.S. performed numerical experiments, collected and analyzed data and drafted the manuscript; S.S., M.S.B., D.Y.L., and Y.K.K. collected data and critically revised the manuscript; D.S.L. critically revised the manuscript; J.S.L. provided initial idea, verified experiments, and critically revised the manuscript. All authors approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-020-80930-w>.

Correspondence and requests for materials should be addressed to J.S.L.

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

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021