RESEARCH PAPER

T1-weighted and T2-weighted MRI image synthesis with convolutional generative adversarial networks

Daisuke Kawahara[1], Yasushi Nagata[1, 2]

[1]Department of Radiation Oncology, Institute of Biomedical and Health Sciences, Hiroshima University, Hiroshima, Japan
[2]Hiroshima High-Precision Radiotherapy Cancer Center, Hiroshima, Japan

ABSTRACT

Background: The objective of this study was to propose an optimal input image quality for a conditional generative adversarial network (GAN) in T1-weighted and T2-weighted magnetic resonance imaging (MRI) images.

Materials and methods: A total of 2,024 images scanned from 2017 to 2018 in 104 patients were used. The prediction framework of T1-weighted to T2-weighted MRI images and T2-weighted to T1-weighted MRI images were created with GAN. Two image sizes (512 × 512 and 256 × 256) and two grayscale level conversion method (simple and adaptive) were used for the input images. The images were converted from 16-bit to 8-bit by dividing with 256 levels in a simple conversion method. For the adaptive conversion method, the unused levels were eliminated in 16-bit images, which were converted to 8-bit images by dividing with the value obtained after dividing the maximum pixel value with 256.

Results: The relative mean absolute error (rMAE) was 0.15 for T1-weighted to T2-weighted MRI images and 0.17 for T2-weighted to T1-weighted MRI images with an adaptive conversion method, which was the smallest. Moreover, the adaptive conversion method has a smallest mean square error (rMSE) and root mean square error (rRMSE), and the largest peak signal-to-noise ratio (PSNR) and mutual information (MI). The computation time depended on the image size.

Conclusions: Input resolution and image size affect the accuracy of prediction. The proposed model and approach of prediction framework can help improve the versatility and quality of multi-contrast MRI tests without the need for prolonged examinations.

Key words: convolutional generative adversarial networks; image synthesis; MRI

(Rep Pract Oncol Radiother 2021;26(1):35–42)

Introduction

Magnetic resonance imaging (MRI) has been widely used; it is non-intrusive and it can provide functional information. There are different contrast images in magnetic resonance MRI types. T1-weighted MRI enhances the signal of the fatty tissue and suppresses the signal of the water. T2-weighted MRI enhances the signal of the water. Consideration of all the information provided by these modalities is conducive to MRI image analysis and diagnosis. On the other hand, the disadvantage of the MRI is that it takes a long time to scan, which might cause a motion artifact [1, 2].

Deep learning has been used for the image-to-image translation networks for image prediction problems such as mapping images to edges and segments, and mapping labels to realistic images. Han performed image-to-image translation between T1-weighted MRI and computed tomography (CT)

images by convolutional neural networks (CNNs) [3]. Convolutional neural network is one of the most typical structures of deep learning. Convolutional neural network can take in an input image and differentiate one from the other by assigning importance (learnable weights and biases) to various aspects/objects in the image. Therefore, CNNs have been widely used in medical image analyses, such as detection and classification [4, 5]. On the other hand, CNN consists of convolutional, pooling, and fully connected layers, adding more layers in CNNs may simply lead to a degradation problem or, in other words, it has been shown that the accuracy degrades rapidly by increasing the network depth [6]. Recently, a generative adversarial network (GAN) has been used that has two networks which is simultaneously CNN [6].

The GAN has a generator that synthesizes data and discriminator that distinguishes between the real and synthesized data. These models can synthesize more realistic images by incorporating the conventional synthesis error and an adversarial loss. Nie et al. proposed the prediction framework from MRI image to CT image with GAN [8–10].

Moreover, Yang et al. proposed the cross-modality generation framework in MRI with GAN [11]. They resized the MRI images to a resolution of 256 × 256 pixels. Also, they used the input image which scaled from 8-bit images (0–255) to a range of –1 to 1. It might reduce the features of the image synthesis by downsizing the image size or reducing the range of the grayscale level.

The current study proposed the prediction framework with GAN which had two CNNs: generator synthesis the T1-weighted MRI image from T2-weighted MRI image or T2-weighted MRI image from T1-weighted MRI image, and the discriminator evaluates them by comparing them with real images. Moreover, the appropriate input image parameters for the generated image are proposed.

Material and methods

Data acquisition

Images from 104 patients were retrospectively analyzed under institutional-review-board-approved study. The MRI image for each patient was acquired with a 3.0 T Philips Ingenia scanner using T1-weighted and T2-weighted MRI images.

Table 1. Magnetic resonance imaging (MRI) scan parameters

| Parameter	T1	T2
TE	23 ms	100 ms
TR	2100 ms	3000 ms
TI	1000 ms	–
Acquired matrix	288 × 197	512 × 336
FOV	220 mm	220 mm
Voxel size	0.76 × 0.76 × 1.12 mm	0.43 × 0.43 × 0.65 mm
Scan time [min]	2.12	2.36

TE — time to Echo; TR — repetition time; TI — inversion time; FOV — field-of-view

The scan parameters for each image are shown in Table 1.

GAN model

In a CNN, local neighborhood pooling operations and trainable filters are applied with the raw input image [12, 13]. Convolutional neural network is superior in image classification and visual object recognition tasks when trained with appropriate regularization [14]. Generative adversarial networks proved superior in image generation [15]. Generative adversarial networks use two different networks of a generator and discriminator networks.

The generator and discriminator networks are trained simultaneously, while the generator attempts to produce realistic images that confuse the discriminator. In the current study, the GAN framework shown in Figure 1 was designed. An approach is proposed for predicting a framework that is based on a conditional GAN with a generator that estimates the MRI image and discriminator that distinguishes the real or synthesis MRI images. Hyperparameter optimization was performed within the training data set, and the evaluation for each algorithm was performed with the test set. The channels in a past study are commonly used as image red–green–blue (RGB) channels in most neural networks [16]. Figure 2 shows the U-net which was used in the generator. Two types of architecture of the U-net were used according to the input images of sizes 512 × 512 and 256 × 256. The transformation of the RGB channel from 16-bit grayscale used a linear interpolation at the same location in the MRI image. However, there are many unused grayscale levels. In the current study, 8-bit RGB
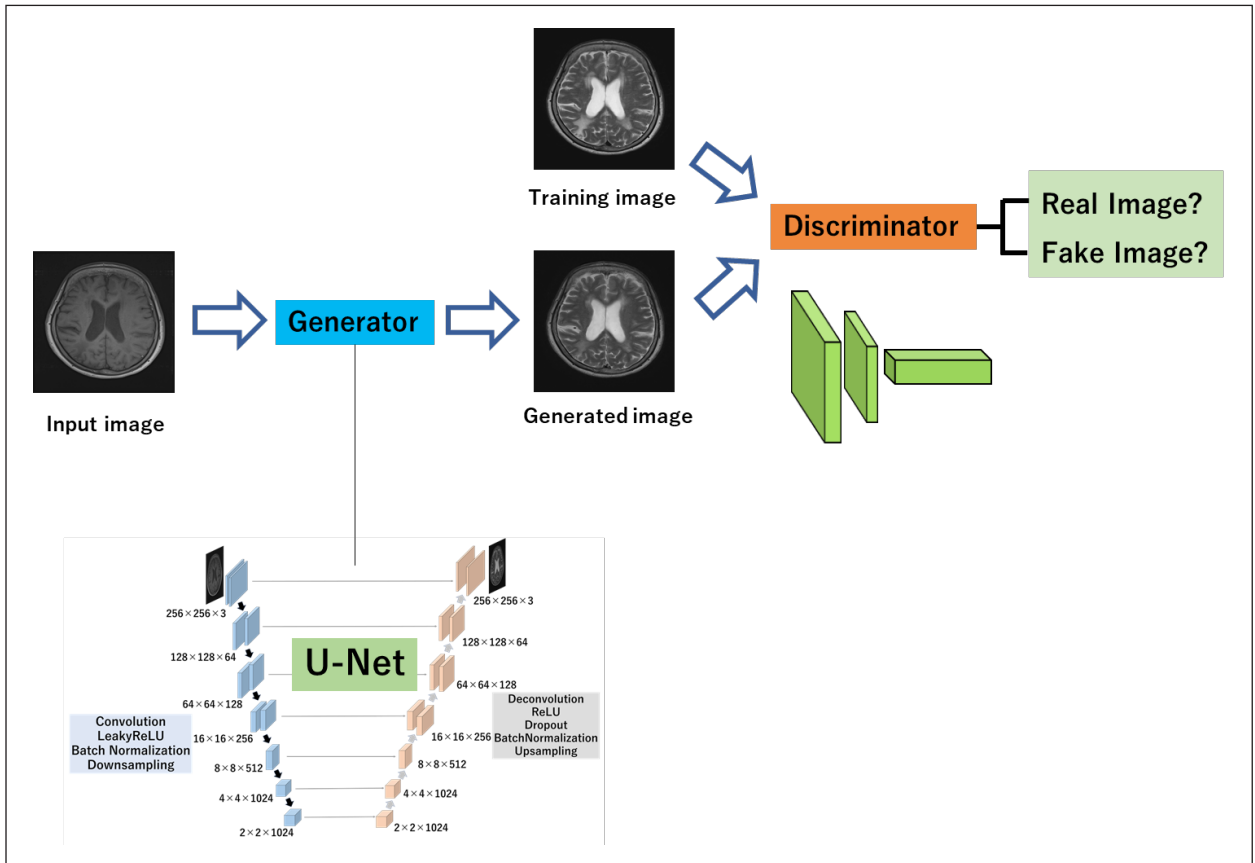
Figure 1. A generative adversarial network (GAN) framework. Given an input MRI, the image in a target contrast with the same anatomy is created by Generator. Discriminator discriminates between real and synthesized MRI (T1-weighted and T2-weighted) images. The subnetworks of Generator and Discriminator are trained simultaneously. Generator minimizes an adversarial loss function and a pixel-wise. Discriminator aims to maximize the adversarial loss function
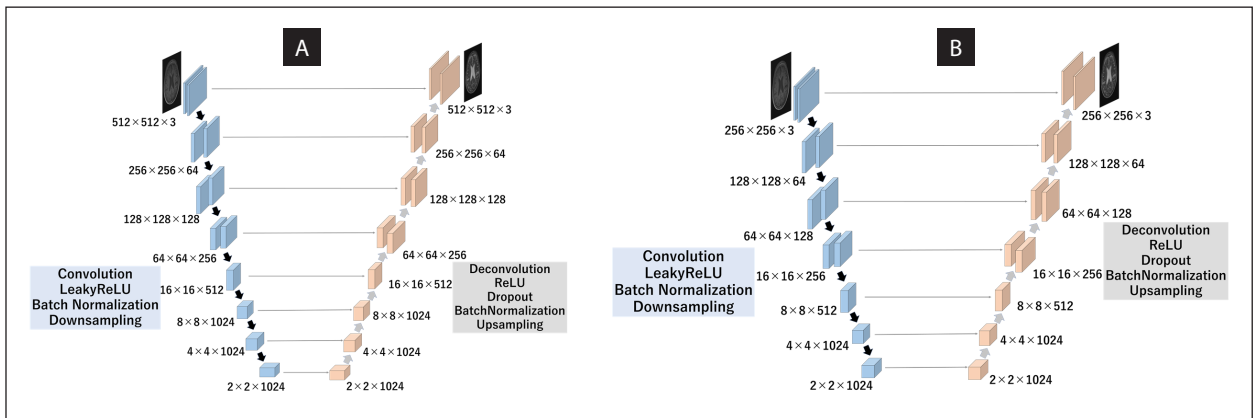
Figure 2. Overall architecture of the U-net used in Generator of the proposed DCNN model. The current study used two types of image sizes 512 × 512 (**A**) and 256 × 256 (**B**)

PNG images were converted from 16-bit DICOM images as shown in Figure 3 with the following two methods. One is the simple conversion method in which a 16-bit (0–65,536) image is converted to an 8-bit (0–255) image by dividing with 256 levels. The other is the adaptive conversion method, in which the unused levels are first eliminated in the 16-bit (0–65,536) images. Then, conversion to 8-bit
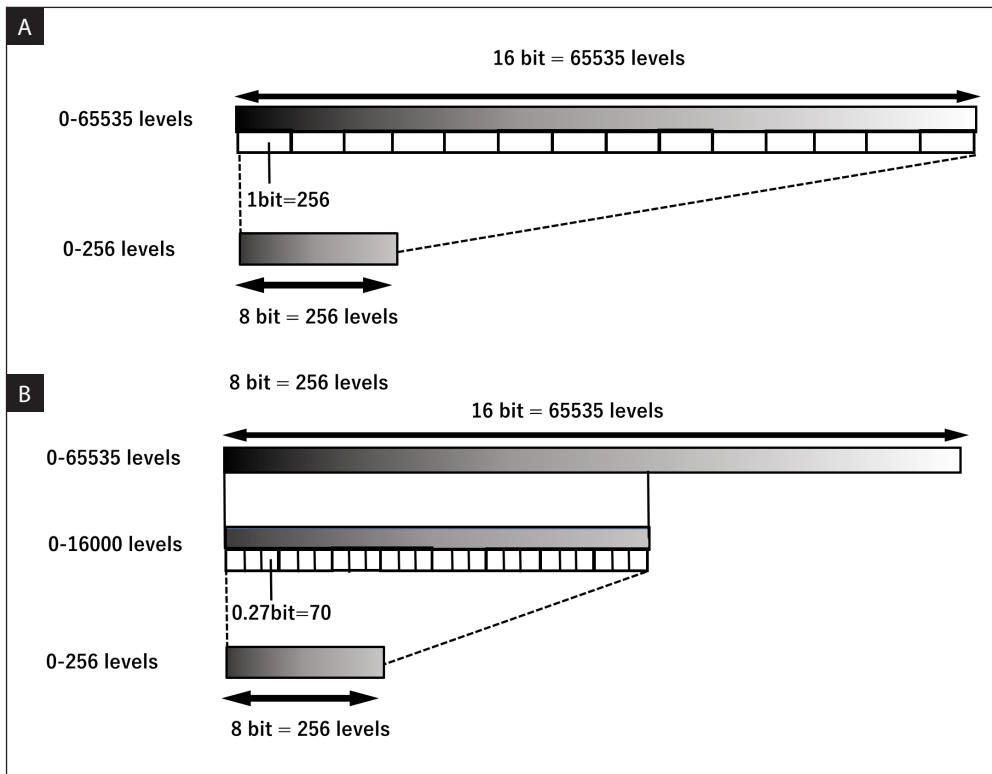
Figure 3. Grayscale conversion methods: **A**. Simple conversion method in which a 16-bit (0–65,536) image was converted to an 8-bit (0–255) one by dividing by 256 levels. **B**. Adaptive conversion method, in which the unused levels were eliminated in the 16-bit (0–65,536) image first and, then, conversion to an 8-bit (0–255) one was performed by dividing by 0.27 bit, in which the maximum pixel value was divided by 256

(0–255) images is performed by dividing with 0.27 bits, which is the value obtained by dividing the maximum pixel value of T1-weighted and T2-weighted MRI images by 256. The dataset of 2,024 images consisted of brain T1-weighted and T2-weighted MRI images from 104 patients. The data were split into two sets: the training used 1,514 images of 70 patients and the model testing used 510 images (34 patients).

The proposed models were implemented using the Ubuntu system (16.04) with TensorFlow packages (V1.7.0, CUDA 9.0, Python 2.7). Except for the batch size, all models were trained with identical hyperparameters and instance normalization. A minibatch of 2D MRI images was selected from the training set at each iteration randomly. Three hundred epochs were used to run on an 11-GB NVIDIA GeForce GTX 1080 GPU.

Evaluation

The prediction accuracy of the model for synthesized MRI and real MRI images was evaluated for the following five metrics: the relative mean absolute error (rMAE), relative mean square error (rMSE), relative root mean square error (rRMSE), peak signal-to-noise ratio (PSNR), and mutual information (MI). The *rMAE* was defined as

$$rMAE = \frac{1}{n_x n_y} \sum_{i,j}^{n_x n_y} \frac{|r(i,j) - s(i,j)|}{r(i,j)}$$

Here $r(i,j)$ is the value of pixel *(i,j)* in the real MRI image, $s(i,j)$ is the value of pixel *(i,j)* in the synthesized MRI image, and $n_x n_y$ is the total number of pixels. *rMSE* is the maximum signal intensity possible and the mean square error (or difference) of the image, which can be presented as

$$rMSE = \frac{1}{n_x n_y} \sum_{i,j}^{n_x n_y} \left(\frac{r(i,j) - s(i,j)}{r(i,j)} \right)^2$$

rRMSE is defined as

$$rRMSE = \sqrt{\frac{1}{n_x n_y} \sum_{i,j}^{n_x n_y} \left(\frac{r(i,j) - s(i,j)}{r(i,j)}\right)^2}$$

PSNR can be calculated as

$$rPSNR_{GL} = 10 \times log_{10}\left(\frac{(d)^2}{rMSE}\right)$$

Here d is a dynamic range. MI is used as a cross-modality similarity measure [13]. It is calculated as

$$I(r:s) = \sum_{m \in I_r} \sum_{n \in I_s} p(m,n) log\left(\frac{p(m,n)}{p(m)p(n)}\right)$$

Here m and n are the intensities in the synthesized MRI image I_r and real MRI image I_s, respectively. p(m) and p(n) represent the marginal probability distributions of m and n. p(m, n) is the joint probability distribution of I_r and I_s. Statistical Analysis was performed using paired t-tests.

Results

Figure 4 shows samples of cross-modality generation results for the synthesized T2-weighted MRI
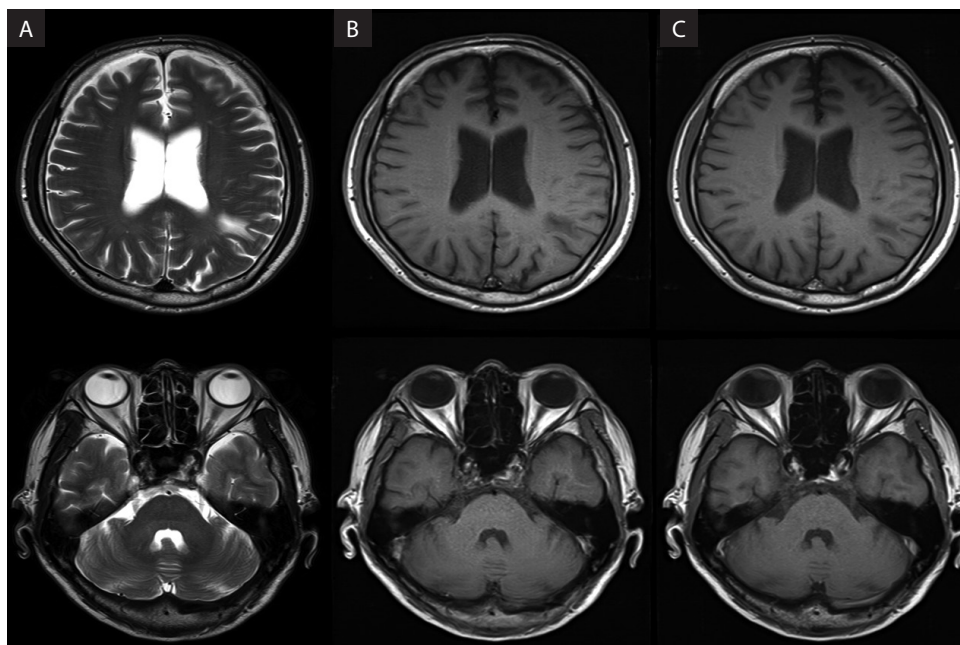
images with input image size of 512 × 512 in the adaptive conversion method. The proposed prediction framework learned anatomical structures such as gray matter, white matter, optic nerve, and brain stem in T1-weighted MRI images. The synthesized T2-weighted MRI images can be reproduced with similar pixel intensity in these structures for real T2-weighted MRI images. Table 2 shows the average rMAE, rMSE, rRMSE, PSNR, and MI computed from the bone and tissue regions for synthesized T1-weighted and T2-weighted MRI images. The adaptive conversion method significantly decreased the rMAE, rMSE, and rRMSE (paired t-test; p < 0.01). The PSNR and MI were significantly larger for the adaptive conversion method. For the comparison of the simple conversion methods with the input image sizes of 512 × 512 and 256 × 256, the rMAE, rMSE, and rRMSE with the high-resolution image were smaller than that with the low-resolution image (paired t-test; p < 0.01), and the PSNR and MI with the high-resolution image were larger than that with the low-resolution image (paired t-test; p < 0.01). The computation time means for the training to create the prediction model. However, the computation time of the high-resolution input image is approximately 3.5 times longer than the time taken by the lower-resolution input image. In the case of the same resolution input images, there
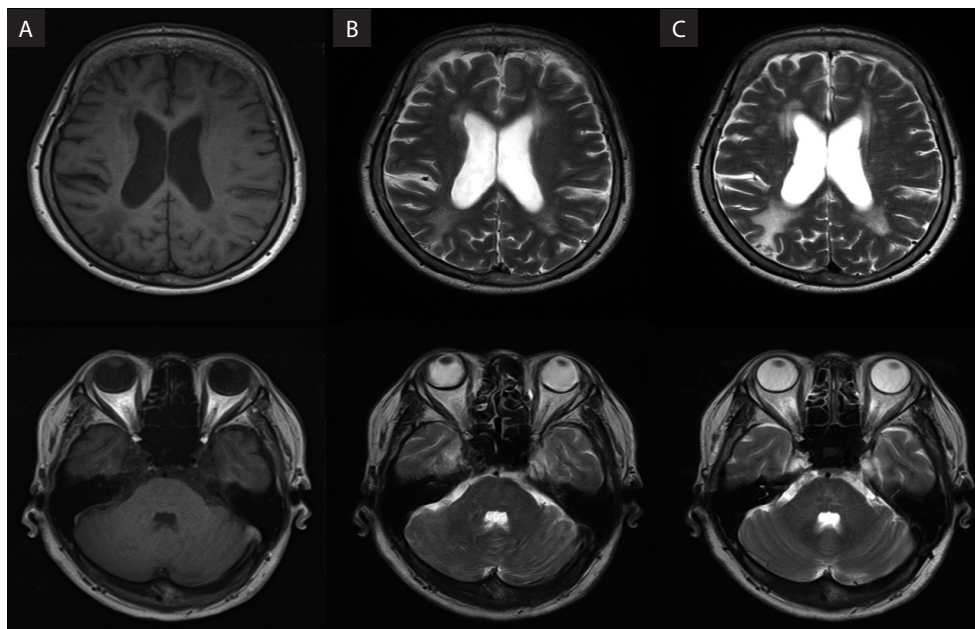
Figure 4. Samples of cross-modality generation results. **A.** T1-weighted MRI image. **B.** Synthesized T2-weighted MRI image. **C.** Real T2-weighted MRI image

Table 2. Average relative mean absolute error (rMAE), smallest mean square error (rMSE), root mean square error (rRMSE), the largest peak signal-to-noise ratio (PSNR), and mutual information (MI) computed from the synthesized T1-weighted MRI image from the T2-weighted MRI images. The time to create the training model with 1,514 images of 70 patients is defined as the computing time

Method	rMAE	rMSE	rRMSE	PSNR	MI	Computing time [min]
Simple conversion (512 × 512)	0.19	0.10	0.32	44.84	1.12	2,794
Simple conversion (256 × 256)	0.28	0.15	0.38	42.26	1.04	785
Adaptive conversion (512 × 512)	0.15	0.07	0.24	58.53	1.30	2,780

Figure 5. Samples of cross-modality generation results. **A.** T2-weighted MRI image. **B.** Synthesized T1-weighted MRI image. **C.** Real T1-weighted MRI image

was a small difference in computation time between the adaptive and simple conversion methods.

Figure 5 shows samples of cross-modality generation results for the synthesized T1-weighted MRI images with input image size of 512 × 512 in the adaptive conversion method. The synthesized T1-weighted MRI images can be reproduced with similar pixel intensity in gray matter,

white matter, optic nerve, and brain stem for real T1-weighted MRI images. Table 3 shows the average rMAE, rMSE, rRMSE, PSNR, and MI computed from the bone and tissue regions for synthesized T1-weighted MRI images. The adaptive conversion method significantly decreased the rMAE, rMSE, and rRMSE (paired t-test; p < 0.01). The PSNR and MI were significantly larger for the adaptive

Table 3. Average relative mean absolute error (rMAE), smallest mean square error (rMSE), root mean square error (rRMSE), the largest peak signal-to-noise ratio (PSNR), and mutual information (MI) computed from the synthesized T2-weighted MRI image from the T1-weighted MRI images. The time to create the training model with 1,514 images of 70 patients is defined as the computing time

Method	rMAE	rMSE	rRMSE	PSNR	MI	Computing time [min]
Simple conversion (512 × 512)	0.20	0.15	0.40	42.54	0.99	2,690
Simple conversion (256 × 256)	0.24	0.19	0.47	41.80	0.95	792
Adaptive conversion (512 × 512)	0.17	0.1	0.31	56.90	1.15	2,672

conversion method (paired t-test; p < 0.01). In the comparison of the simple conversion methods with input image sizes of 512 × 512 and 256 × 256, the high-resolution input image significantly decreased the rMAE, rMSE and rRMSE (paired t-test; p < 0.01). The PSNR and MI were significantly larger with the high-resolution input image (paired t-test; p < 0.01). However, the computation time of the high-resolution input image is approximately 3.3 times longer than the lower-resolution input image. In the case of the same resolution input images, there was a small difference in computation time between the adaptive and simple conversion methods.

The times to create the synthesized MRI images using all the trained models were approximately 28–30 s for the conversion of both T1-weighted to T2-weighted MRI images and T2-weighted to T1-weighted MRI images.

Discussion

The current study used the conditional GAN with a generator and a discriminator for generating brain MRIs based on GANs from T1-weighted or T2-weighted images and evaluated the appropriate input image parameters. In a prior study, Jog et al. proposed the random forest method in multicontrast MRI image synthesis [17]. The PSNR was 24.7–28.0 in the synthesis from T1-weighted to T2-weighted MRI images and 22.3–23.4 in the synthesis from T2-weighted to T1-weighted MRI images, which were lower than the current study. The prediction by random forest method is averaged during the process of image synthesis. It may lead to a loss of high spatial frequency information. The GAN model proposed in the current study trained the image generation and distinguishing the image of the real or synthesized images, simultaneously. By incorporating the adversarial loss, the detail of the texture information in the target contrast is captured, thereby making higher synthesis quality possible by comparison with absolute error loss or typical squared.

Yang et al. predicted the T2-weighted MRI images from the T1-weighted MRI images with a conditional GAN [11]. The current study increased the convolutional layers and expanded the range of the grayscale in the images. The current study had a higher PSNR and MI for both image synthesis

of T1-weighted to T2-weighted MRI images and T2-weighted to T1-weighted MRI images.

The input image matrix size affected the synthesized image quality. The high-resolution input image increased the accuracy of the image synthesized. The features around the tissue or lesion are seen better in the high-resolution image. Although the computation time to create the training model is long for the larger input size, there is little difference in the generating time for the synthesized MRI image after training. The smaller input size images should be used when evaluating the optimal training model. To consider the introduction of the clinical model, the training model should be created without resizing the scanned image size. The adaptive conversion method additionally increased the accuracy of the synthesized image quality. The quantity of the information using the adaptive conversion method can increase more than three times compared with the simple conversion method. The sufficient image level in learning contributes to the increase of the tissue contrast. The adaptive conversion method can generate higher-quality results without increasing the time to create the training model.

In clinical applications, MRI is widely used because of the diversity of contrast in soft tissues. Each MRI pulse sequence can generate distinct contrasts within the same anatomy. In turn, the available diagnostic information in clinical practice is increased by using multi-contrast MRI images acquired in the same subject. However, the multicontrast MRI image requires a long acquisition time. Moreover, a past study reported that a subset of the acquired contrasts was corrupted by artifacts or excessive noise artifacts occasionally [18]. The proposed model of image synthesis can obtain a multi-contrast MRI image using one sequence in a few seconds for one patient. It leads to a reduction of the moving artifact or any other artifact. Thus, the method can improve the scan time dramatically and improve analysis tasks, such as segmentation and registration. For radiation diagnosis, the multi-contrast MRI image should assist in lesion detectability. Further study is needed to evaluate the detectability of the lesion.

Conclusion

Input resolution and image size affect the accuracy of prediction. The proposed adaptive conversion

model and approach of prediction framework can help reduce the scan time and improve the versatility and quality of multi-contrast MRI tests.

Ethnical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

Informed consent was obtained from all individual participants included in the study.

Research involving human participants and/or animals

This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest

None declared.

Financial disclosure

None declared.

Acknowledgements

None declared.

Data sharing

All data available.

Advances in knowledge

We proposed the optimal input image size and the conversion method of the gradient for the image synthesis model in T1-weighted and T2-weighted MRI images.

References

1. Rzedzian R, Chapman B, Mansfield P, et al. Real-time nuclear magnetic resonance clinical imaging in paediatrics. Pediatr Radiol. 1986; 2(8362): 1281–1282, doi: 10.1016/s0140-6736(83)91153-4, indexed in Pubmed: 6139622.
2. Tsao J. Ultrafast imaging: principles, pitfalls, solutions, and applications. J Magn Reson Imaging. 2010; 32(2): 252–266, doi: 10.1002/jmri.22239, indexed in Pubmed: 20677249.
3. Han X. MR-based synthetic CT generation using a deep convolutional neural network method. Med Phys. 2017; 44(4): 1408–1419, doi: 10.1002/mp.12155, indexed in Pubmed: 28192624.
4. Shen W, Zhou Mu, Yang F, et al. Multi-scale Convolutional Neural Networks for Lung Nodule Classification. Inf Process Med Imaging. 2015; 24: 588–599, doi: 10.1007/978-3-319-19992-4_46, indexed in Pubmed: 26221705.
5. Kooi T, Litjens G, van Ginneken B, et al. Large scale deep learning for computer aided detection of mammographic lesions. Med Image Anal. 2017; 35: 303–312, doi: 10.1016/j.media.2016.07.007, indexed in Pubmed: 27497072.
6. He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016, doi: 10.1109/cvpr.2016.90.
7. Isola P, Zhu JY, Zhou T, et al. Image-to-Image Translation with Conditional Adversarial Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017, doi: 10.1109/cvpr.2017.632.
8. Nie D, Trullo R, Lian J, et al. Medical Image Synthesis with Context-Aware Generative Adversarial Networks. Med Image Comput Comput Assist Interv. 2017; 10435: 417–425, doi: 10.1007/978-3-319-66179-7_48, indexed in Pubmed: 30009283.
9. Nie D, Trullo R, Lian J, et al. Medical Image Synthesis with Deep Convolutional Adversarial Networks. IEEE Trans Biomed Eng. 2018; 65(12): 2720–2730, doi: 10.1109/TBME.2018.2814538, indexed in Pubmed: 29993445.
10. Emami H, Dong M, Nejad-Davarani SP, et al. Generating synthetic CTs from magnetic resonance images using generative adversarial networks. Med Phys. 2018 [Epub ahead of print], doi: 10.1002/mp.13047, indexed in Pubmed: 29901223.
11. Yang Q, Li N, Zhaob Z. MRI Cross-Modality NeuroImage-to-NeuroImage Translation. arXiv. 2018; 1801: 06940v2.
12. Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. JMLR Workshop and Conference Proceedings. 2010; 9: 249–256.
13. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521(7553): 436–444, doi: 10.1038/nature14539, indexed in Pubmed: 26017442.
14. LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. Proceedings of the IEEE. 1998; 86(11): 2278–2324, doi: 10.1109/5.726791.
15. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. Advances in Neural Information Processing Systems. 2014; 27: 2672–2680.
16. Zhou X, Takayama R, Wang S, et al. Deep learning of the sectional appearances of 3D CT images for anatomical structure segmentation based on an FCN voting method. Med Phys. 2017; 44(10): 5221–5233, doi: 10.1002/mp.12480, indexed in Pubmed: 28730602.
17. Jog A, Roy S, Carass A, et al. Magnetic resonance image synthesis through patch regression. Proc IEEE Int Symp Biomed Imaging. 2013; 2013: 350–353, doi: 10.1109/ISBI.2013.6556484, indexed in Pubmed: 24443686.
18. Krupa K, Bekiesińska-Figatowska M. Artifacts in magnetic resonance imaging. Pol J Radiol. 2015; 80: 93–106, doi: 10.12659/PJR.892628, indexed in Pubmed: 25745524.