Medical Image Synthesis with Deep Convolutional Adversarial Networks

Dong Nie, Roger Trullo, Jun Lian, Li Wang, Caroline Petitjean, Su Ruan, Qian Wang, and Dinggang Shen

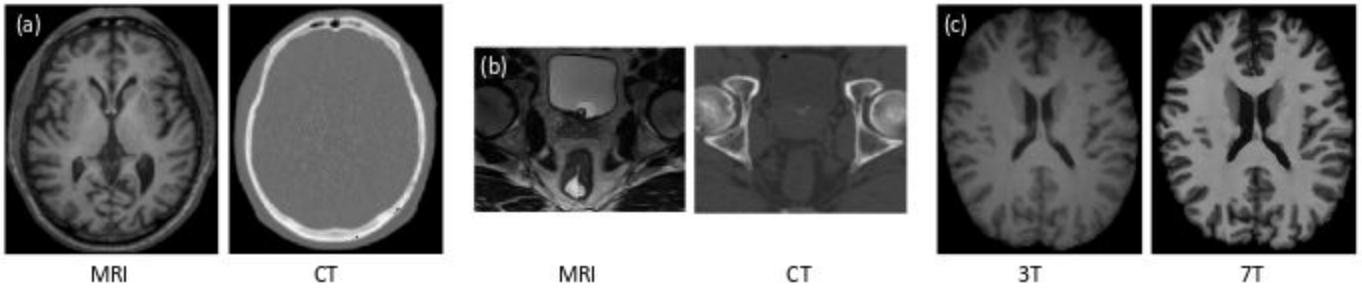

Fig. 1. Three pairs of corresponding source (left) and target (right) images from the same subjects. (a) shows a pair of MRI/CT brain images; (b) shows a pair of MRI/CT pelvic images; (c) shows a pair of 3T/7T brain MRI.

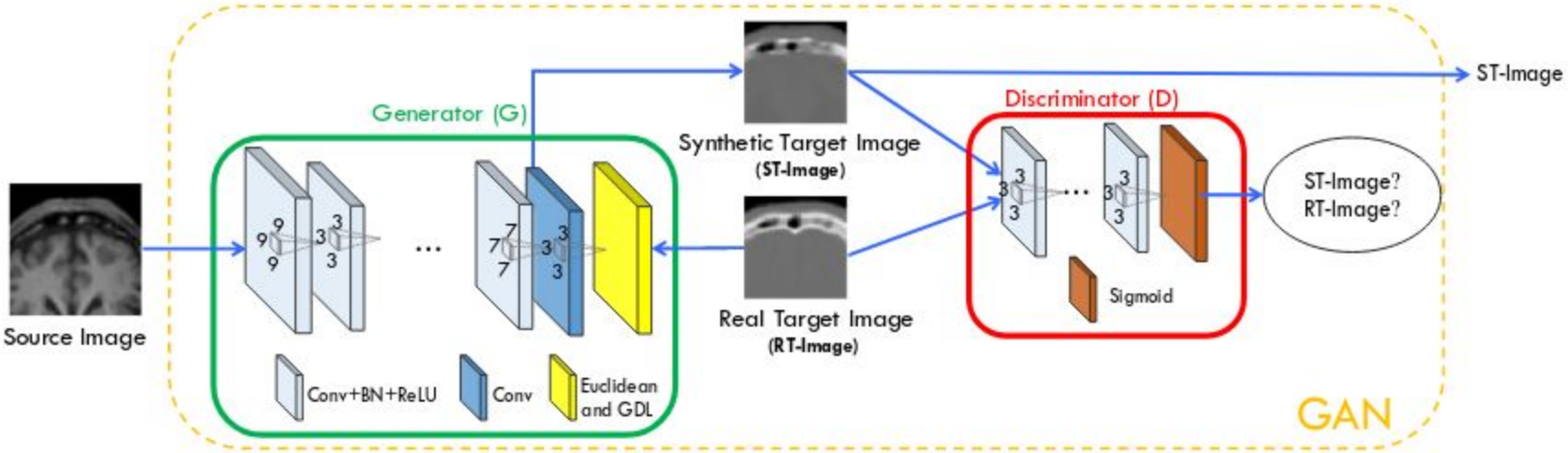

Fig. 2. Architecture used in the deep convolutional adversarial setting for estimation of the synthetic target image.

Architecture Details

- Generator:
  - 9 layers with convolution, BN, and relu
  - Kernel sizes: 9,3,3,3,9,3,3,7,3
  - Input image size: 32x32x32
  - Number of kernels: 32,32,32,64,64,64,32,32,1
  - Dilation rates for first and last layers: 1; else: 2
- Discriminator:
  - 3 stages of convolution, BN, relu, Max Pooling, followed by 1 convolution and 3 FC layers with 2 relu activations and one sigmoid activation
  - Kernel size for all layers: 3
  - Nodes in FC layers: 512, 128, 1
  - Number of kernels: 32, 64, 128, 256

$$L_G(X,Y) = \|Y - G(X)\|_2^2 \tag{1}$$

$$L_{ADV}(X) = L_{BCE}(D(G(X)),1) \tag{5}$$

$$\begin{aligned} L_{G\_ADV}(X,Y) =& \lambda_1 L_{ADV}(X) \\ &+ \lambda_2 L_G(X,Y) + \lambda_3 L_{GDL}(Y, G(X)) \end{aligned} \tag{4}$$

$$\begin{aligned} L_{GDL}(Y,\hat{Y}) =& \left| |\nabla Y_x| - \left|\nabla \widehat{Y}_x\right| \right|^2 \\ &+ \left| |\nabla Y_y| - \left|\nabla \widehat{Y}_y\right| \right|^2 \\ &+ \left| |\nabla Y_z| - \left|\nabla \widehat{Y}_z\right| \right|^2 \end{aligned} \tag{6}$$

$$L_D(X,Y) = L_{BCE}(D(Y),1) + L_{BCE}(D(G(X)),0) \tag{2}$$

$$L_{BCE}(\widehat{Y},Y) = -\sum_i Y_i \log\left(\widehat{Y}_i\right) + (1 - Y_i)\log\left(1 - \widehat{Y}_i\right) \tag{3}$$

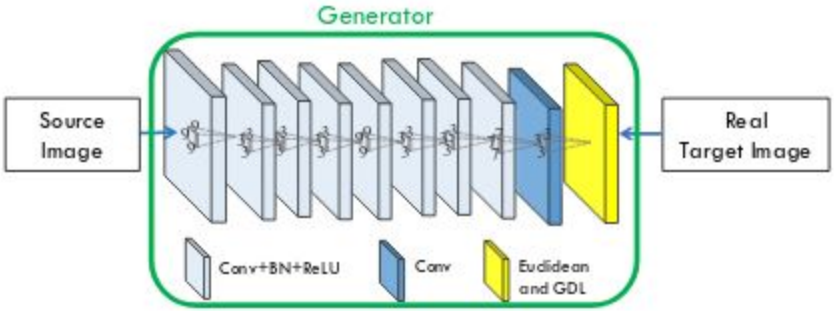

Fig. 3. The 3D FCN architecture for estimating a target image from a source image.

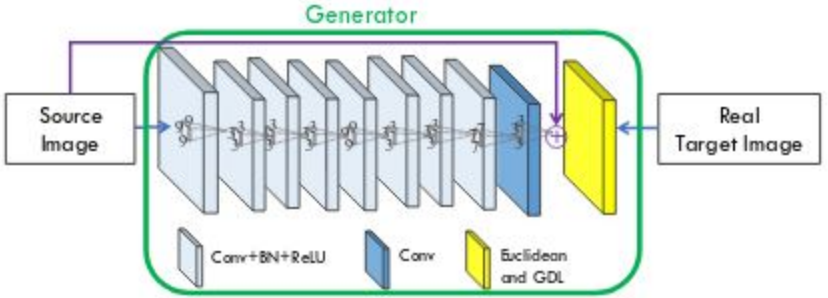

Fig. 4. Generator architecture used in the GAN setting for estimation of synthetic target image. Note the solid-purple-line arrow from source image to the 'plus' sign, which expresses the long-term residual connection.

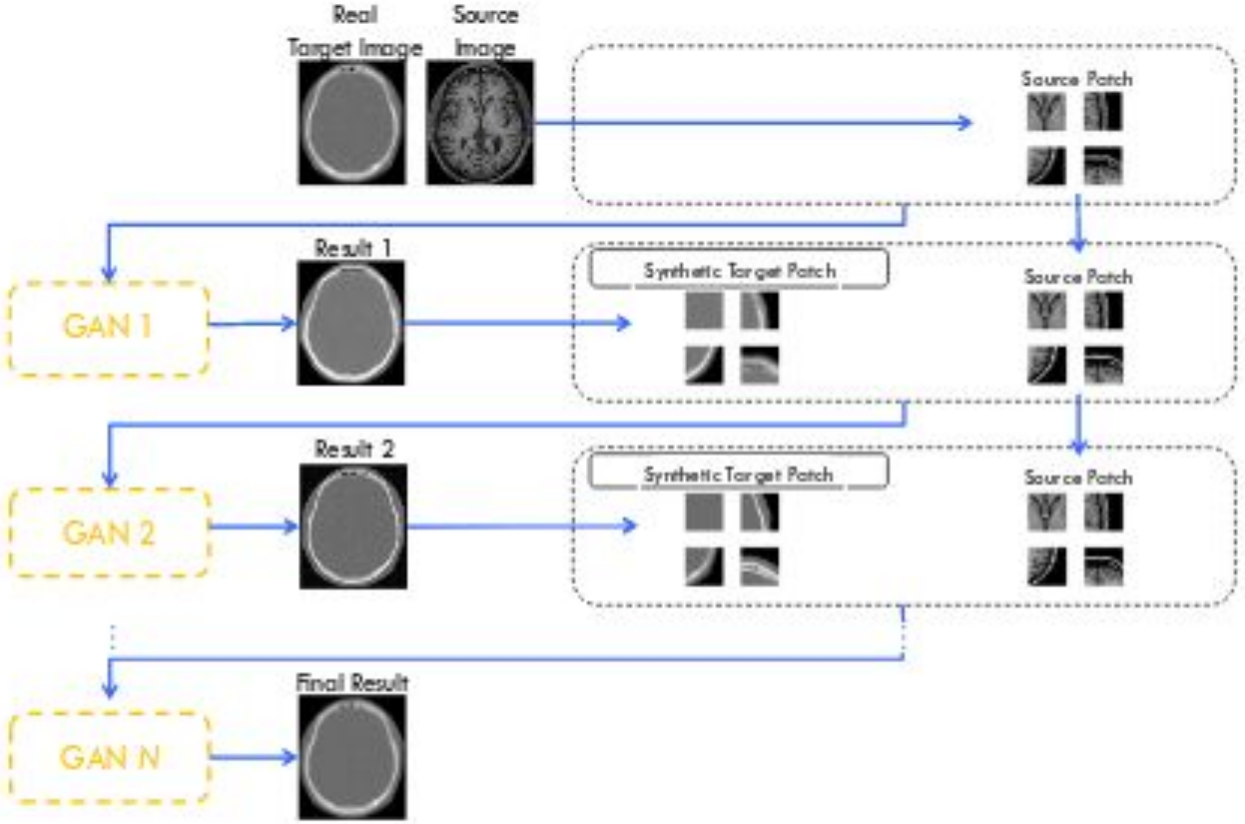

Fig. 5. Proposed architecture for ACM with GAN.

Experimental Controls for CT synthesis

- Atlas-based method: uses atlas-based registration (with Demons) with intensity averaging with 5 atlases (Atlas)
- Sparse Representation based method (SR)
- Structured Random Forest with ACM (SR+)
- Evaluation on their data set was only done using first two methods. They report the results from the paper for SR+

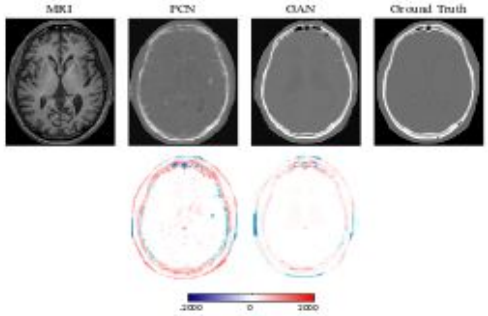

Fig. 6. Visual comparison for impact of adversarial learning. The 1st row shows the synthetic CT by FCN and GAN, and the 2nd row shows the difference map between the synthetic CT and ground truth CT. Note that FCN means the case without adversarial learning, and GAN means the proposed method with adversarial learning.

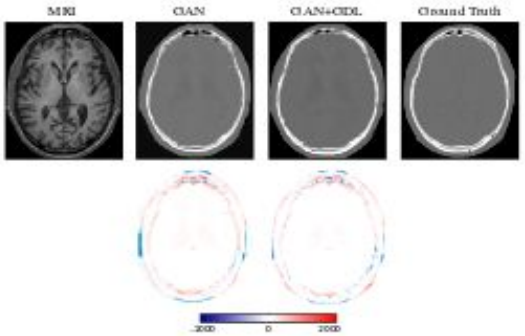

Fig. 7. Visual comparison for impact of using the gradient difference loss (GDL). The 1st row shows the input MRI, two synthetic CT by two different methods, and the ground-truth CT. The 2nd row shows difference maps between each synthetic CT and the ground-truth CT.

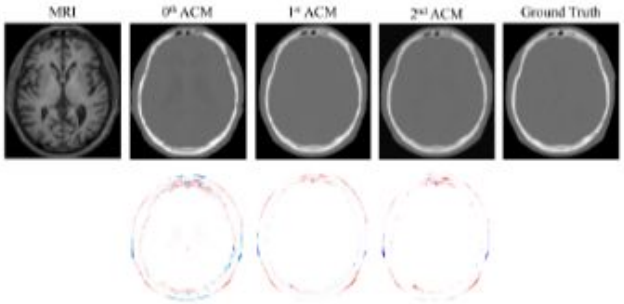

Fig. 10. The 1st row shows visual comparison of MR image, three synthetic CT images by applying 0th, 1st and 2nd iterations of ACM, and the ground-truth CT image for a typical brain case. The 2nd row shows difference maps between each iteratively-estimated CT and the ground-truth CT.

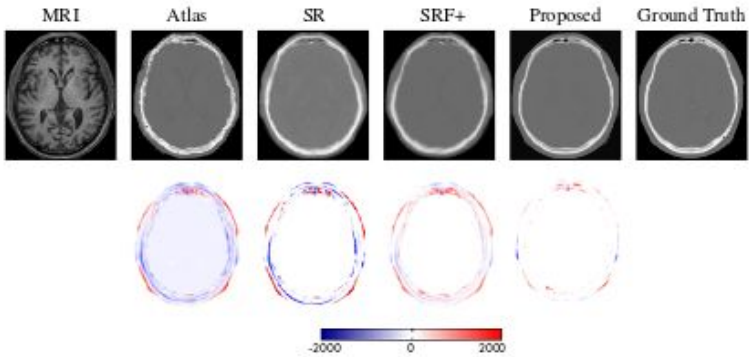

Fig. 11. The 1st row shows visual comparison of the MR image, the four estimated CT images by other three competing methods and our proposed method, and the ground-truth CT for a typical brain case. The 2nd row shows difference maps between each estimated target CT and the ground-truth CT.

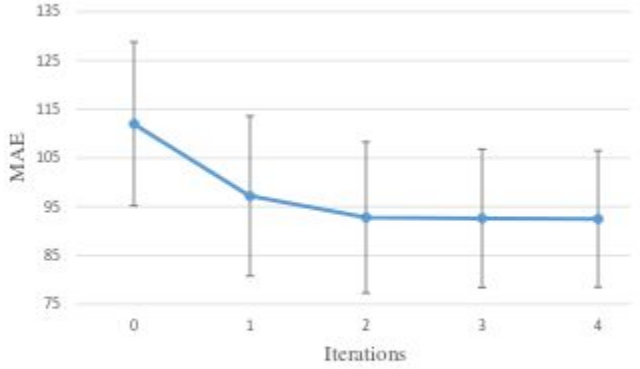

Fig. 9. Performance (MAE) of using ACM on the brain dataset with iterations.

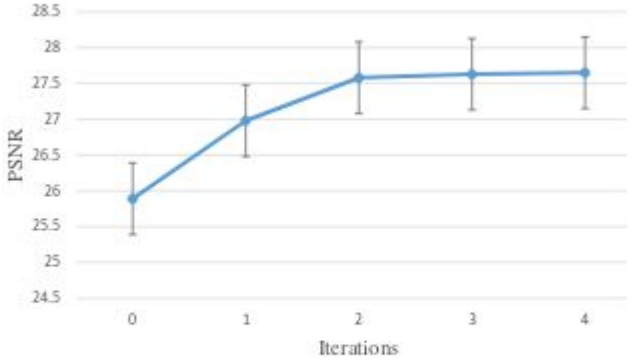

Fig. 8. Performance (PSNR) of using ACM on the brain dataset with iterations.

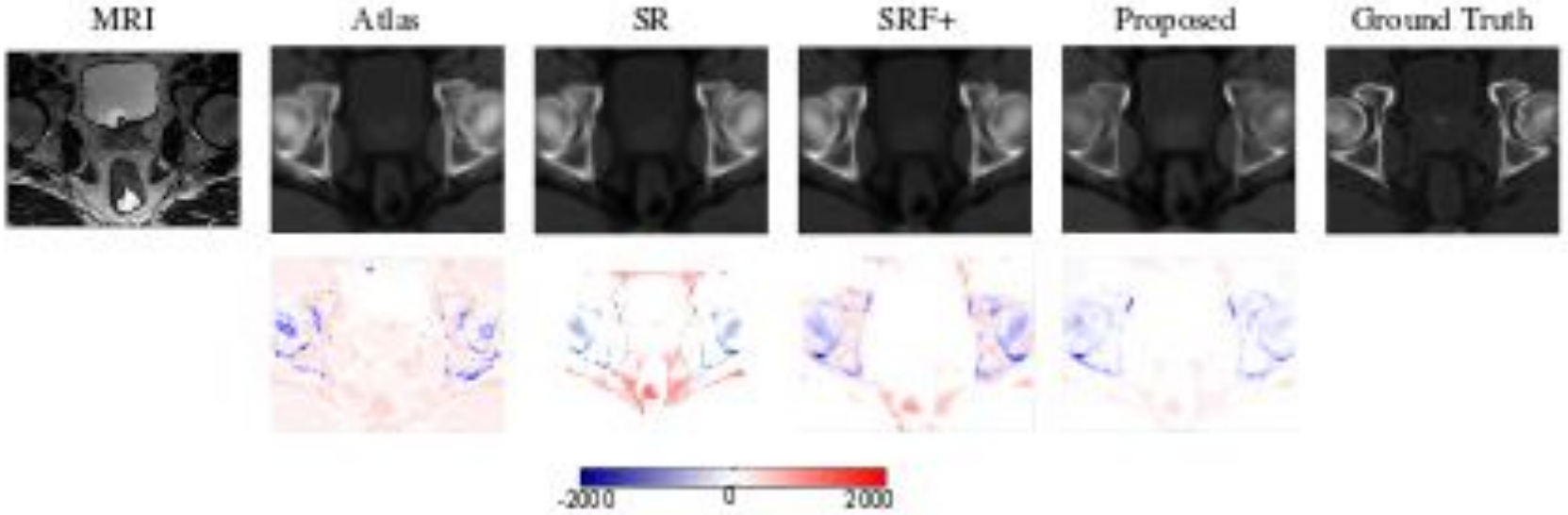

Fig. 12. The 1st row shows visual comparison of the MR image, the estimated CT images by our method and other competing methods, and the ground-truth CT image for the typical pelvic case; The 2nd row shows the difference maps between estimated CT and ground truth CT.

TABLE I

RESULTS SUMMARIZING DIFFERENT EFFECTS OF OUR PROPOSED METHOD ON THE BRAIN DATASET IN TERMS OF PSNR.

Method	No Adv.	Adv.	Adv.+GDL	Proposed
Mean(std)	24.7(1.4)	25.2(1.4)	25.9(1.4)	27.6(1.3)

TABLE II

AVERAGE MAE AND PSNR ON 16 SUBJECTS FROM THE BRAIN DATASET.

Method	MAE		PSNR	
	Mean (std)	Med.	Mean (std)	Med.
Atlas	171.5(35.7)	170.2	20.8(1.6)	20.6
SR	159.8(37.4)	161.1	21.3(1.7)	21.2
SRF+ [18]	99.9(14.2)	97.6	26.3(1.4)	26.3
Proposed	**92.5(13.9)**	**92.1**	**27.6(1.3)**	**27.6**

TABLE III

AVERAGE MAE AND PSNR ON 22 SUBJECTS FROM THE PELVIC DATASET.

Method	MAE		PSNR	
	Mean (std)	Med.	Mean (std)	Med.
Atlas	66.1(6.9)	66.7	29.0(2.1)	29.6
SR	52.1(9.8)	52.3	30.3(2.6)	31.1
SRF+ [18]	48.1(4.6)	48.3	32.1(0.9)	31.8
Proposed	**39.0(4.6)**	**39.1**	**34.1(1.0)**	**34.1**

TABLE IV

P-VALUES BY PERFORMING WILCOXON SIGNED-RANK TEST BETWEEN OUR PROPOSED METHOD AND ALL THE PREVIOUS METHOD FOR BOTH PSNR AND MAE VALUES ON BRAIN AND PELVIC DATASETS.

Method	Brain		Pelvic	
	PSNR	MAE	PSNR	MAE
Atlas	<0.01	<0.01	<0.01	<0.01
SR	<0.01	<0.01	<0.01	<0.01
SRF+ [18]	<0.05	<0.05	<0.01	<0.01

Experimental Controls for 7T MRI synthesis

- Histogram matching: matches intensity distributions of image and target (HM)
- Local Image Similarity: synthesizes target image using multiple atlases with local similarity measures (LIS)
- Multi-level CCA: Hierarchical reconstruction using group sparsity in multi-level framework (M-CCA)
- Convolutional Neural Network: Train CNN to learn mapping from source to target (CNN)

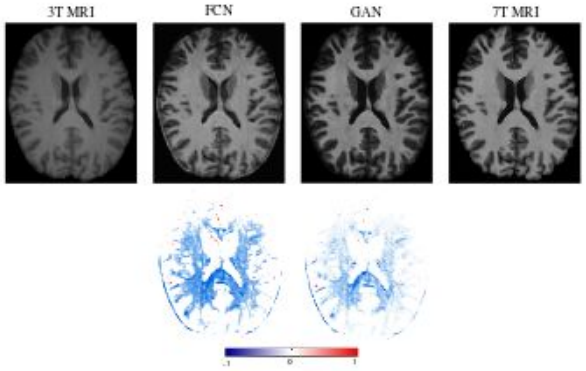

Fig. 13. Visual comparison to demonstrate the impact of using the adversarial learning for the 3T-to-7T dataset. The 1st row shows 3T MRI, two synthetic 7T MRI by two methods, and ground-truth 7T MRI. The 2nd row shows difference maps between each synthetic 7T MRI and ground-truth 7T MRI. Note that FCN means the case without adversarial learning, and GAN means the case with adversarial learning.

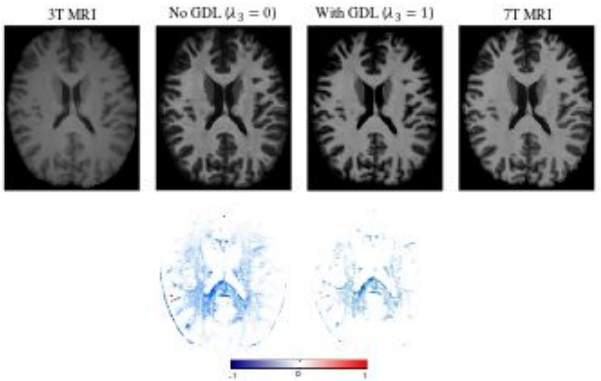

Fig. 14. Visual comparison to demonstrate the impact of using the gradient difference loss. The image obtained via GDL is more realistic and sharper. The 1st row shows 3T MRI, the synthetic 7T MRI by two methods, and ground-truth 7T MRI; the 2nd row shows difference maps between each synthetic 7T MRI and the ground-truth 7T MRI.

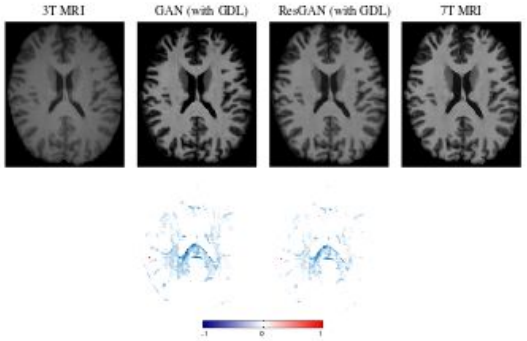

Fig. 15. Visual comparison to demonstrate the impact of using the residual learning. The 1st row shows synthetic 7T MRI, and the 2nd row shows the difference maps between the synthetic 7T MRI and ground truth 7T MRI.

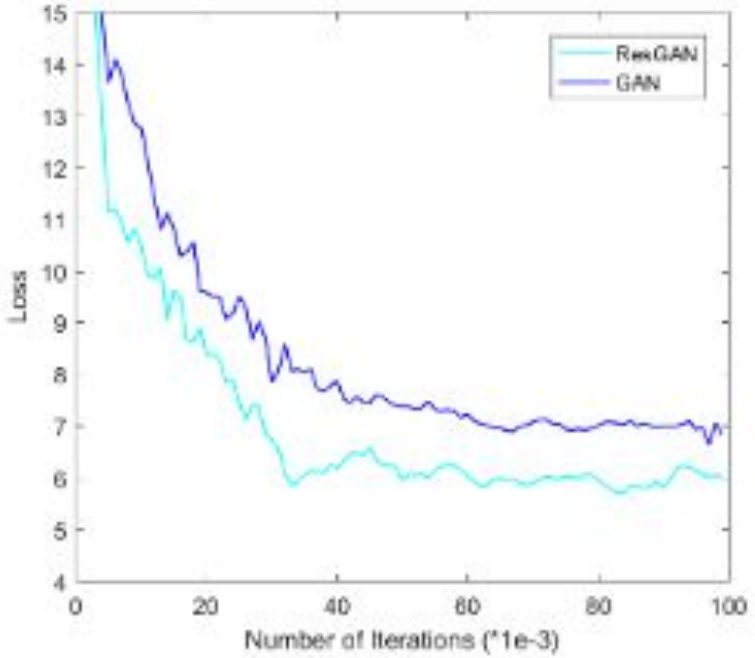

Fig. 16. The mean squared loss (MSE) of the generator in GAN and ResGAN on the testing dataset *with respect to* different training iterations.

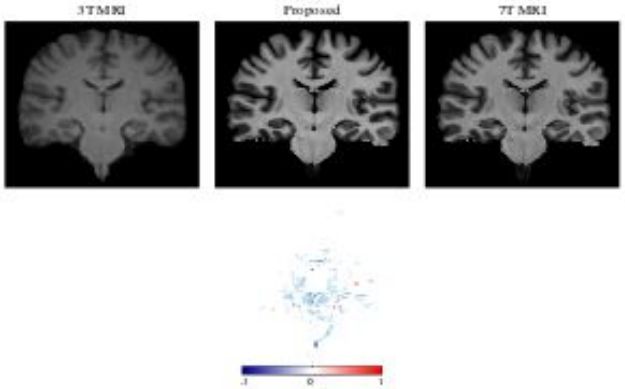

Fig. 17. The visualization for subcortical regions with 3T MRI, the synthetic 7T MRI, and the ground-truth 7T MRI. The 2nd row shows a difference map between the ground-truth 7T MRI and the synthetic 7T MRI.

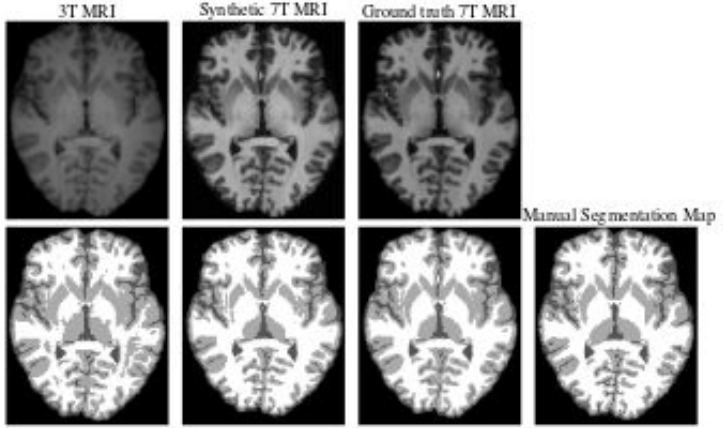

Fig. 18. Visual comparison of segmentation results for a typical subject by using different input data (3T MRI, synthetic 7T MRI, and ground-truth 7T MRI). The 1st row is the MRI, and the 2nd shows the segmented slices as well as the manual segmentation map.

TABLE VI

RESULTS SUMMARIZING DIFFERENT EFFECTS OF OUR PROPOSED METHOD ON THE 3T-TO-7T DATASET IN TERMS OF OF PSNR.

Method	No Adv.	Adv.	Adv.+GDL	ResGAN	ResGAN+ACM
Mean(std)	26.15(1.27)	26.83(1.25)	27.18(1.24)	27.69(1.22)	27.93(1.18)

TABLE V

PERFORMANCE OF SEGMENTATION ON THE MRI DATASET IN TERMS OF DICE INDEX AND ITS CORRESPONDING STANDARD DEVIATION.

Input	WM	GM	CSF
3T MRI	80.35(2.02)	85.49(1.08)	88.75(0.93)
Synthetic 7T MRI	86.84(1.84)	91.68(0.92)	95.96(0.88)
Ground-Truth 7T MRI	87.70(1.76)	92.33(0.86)	96.58(0.90)

TABLE VII

COMPARISON OF THE PERFORMANCES OF DIFFERENT METHODS ON THE 3T-TO-7T DATASET IN TERMS OF PSNR. THE P-VALUES BY PERFORMING WILCOXON SIGNED-RANK TEST BETWEEN OUR PROPOSED METHOD AND ALL OTHER METHODS ARE ALSO REPORTED, AND WE USE "*" TO DENOTE $p<0.01$.

Method	HM	LIS	M-CCA	CNN	Proposed
Mean (std)	21.10(1.44) *	24.33(1.26) *	25.41(1.20) *	26.50(1.22) *	**27.93(1.18)**