<center>**Supplemental Digital Content**</center>

Methods

Network Architecture

The discriminator distinguishes the synthetic 7T images from the real 7T images. It contains a stack of convolutional layers with stride 2 to extract features and reduce size of feature maps. The last layer simply averages out the feature maps to form the decision variable for binary classification. The network architecture is illustrated in Fig. 1 in the Supplementary Material.

Loss Function

Let x denotes a 3T image and y denotes a 7T images, the generator G tries to synthesize a 7T image $G(x)$ that resembles the real 7T image y. The discriminator D tries to distinguish the synthetic image pair $(x, G(x))$ from its corresponding real image pair (x, y). The loss function of the generator G is formulated as follows:

$$L_G = \mathbf{E}_x\left[log\left(1 - D\big(x, G(x)\big)\right)\right] + \lambda_1 \mathbf{E}_{x,y}[\|y - G(x)\|_1] \qquad (1)$$

where $\mathbf{E}[\ ]$ refers to the expectation operator, λ_1 denote the weighting parameter, and θ_D denotes the network parameters of the discriminator D. $\|\ \ \|_1$ denotes a L1-norm that penalty on the pixel-wise intensity difference between the synthetic and real 7T images to ensure intensity similarity. The loss function of the discriminator D is formulated as follows:

$$L_D = -\mathbf{E}_{x,y}[log(D(x, y))] - \mathbf{E}_{x,y}\left[log\left(1 - D\big(x, G(x)\big)\right)\right] \qquad (2)$$

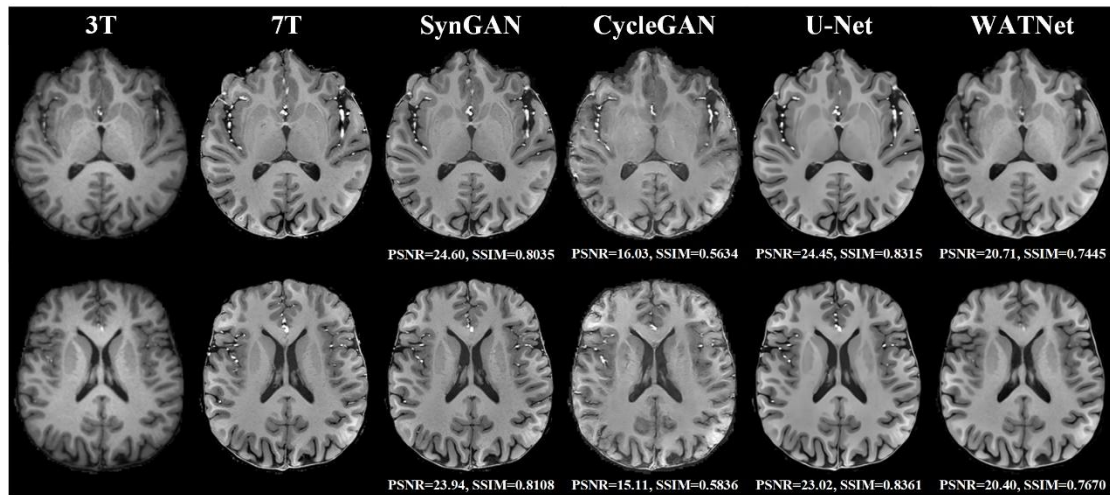Different from G, D tries to assign correct labels to the real and synthetic image pairs. Thus, D and G are following a two-player min-max game in network training.

**Supplementary Table E1: Image quality assessment using the 5-point Likert scale**

| Score	Overall image quality	Artifacts	Sharpness	Contrast	Visualization of vessel
1	Nondiagnostic	Nondiagnostic	Nondiagnostic	Nondiagnostic	Nondiagnostic
2	Substantial deficits in image quality	Substantial impact on diagnosis	Not sharp	Poor contrast	Poor visualization
3	Moderate image quality	Moderate impact on diagnosis	A little sharp	Moderate contrast	Moderate visualization
4	Good image quality	Minimal impact on diagnosis	Moderately sharp	Good contrast	Good visualization
5	Excellent image quality	No artifact	Satisfyingly sharp	Excellent contrast	Excellent visualization
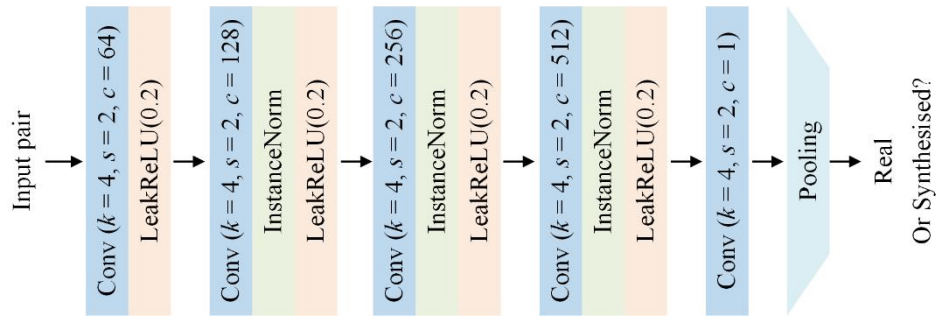
Supplementary Fig. 1: The network architecture of discriminator. The discriminator takes a pair of the 3T image with the real 7T image or synthetic 7T image as input, and outputs a decision variable for binary classification. Conv denotes a convolutional layer, and k, s, c denotes the kernel size, stride, and out channels of the convolutional layer, respectively. InstanceNorm denotes an Instance Normalization layer. LeakReLU(0.2) denotes a LeakyReLU function with negative slope $= 0.2$.

| 3T	7T	SynGAN	CycleGAN	U-Net	WATNet
		PSNR=24.60, SSIM=0.8035	PSNR=16.03, SSIM=0.5634	PSNR=24.45, SSIM=0.8315	PSNR=20.71, SSIM=0.7445
		PSNR=23.94, SSIM=0.8108	PSNR=15.11, SSIM=0.5836	PSNR=23.02, SSIM=0.8361	PSNR=20.40, SSIM=0.7670

Supplementary Fig. 2: Representative synthetic 7T images with different synthesis methods in a 21-year-old healthy participant. It can be observed that SynGAN yields high-quality synthetic images that are significantly closer to the real 7T images. PSNR and SSIM values are shown at the bottom. PSNR = peak signal-to-noise ratio; SSIM = structure similarity.
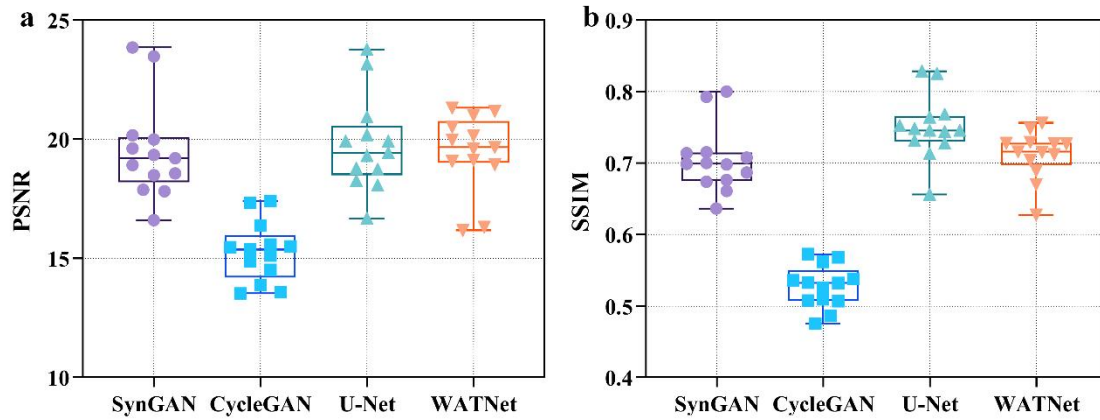
Supplementary Fig. 3: Boxplots of PSNR and SSIM values over the non-enhanced test data for four different 7T image synthesis methods. Each circular dot indicates the PSNR or SSIM value from each subject. The Mean PSNR values for the four methods are: 19.53 ± 2.08 for SynGAN, 15.27 ± 1.25 for CycleGAN, 19.78 ± 1.96 for U-Net, 19.47 ± 1.63 for WATNet. The Mean SSIM values for the three methods are: 0.7047 ± 0.0463 for SynGAN, 0.5271 ± 0.0298 for CycleGAN, 0.7500 ± 0.0441 for U-Net, 0.7118 ± 0.0339 for WATNet. PSNR = peak signal-to-noise ratio; SSIM = structure similarity.
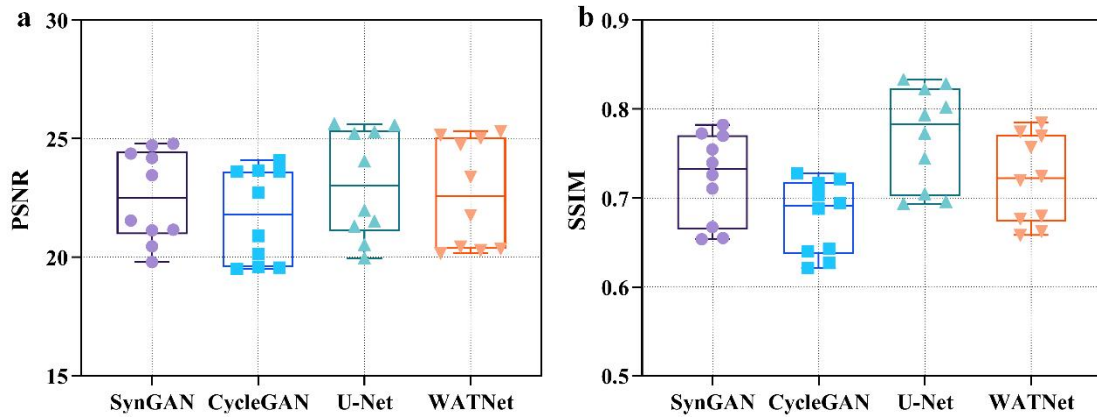
Supplementary Fig. 4: Boxplots of PSNR and SSIM values over the contrast-enhanced test data for four different 7T image synthesis methods. Each circular dot indicates the PSNR or SSIM value from each subject. The Mean PSNR values for the four methods are: 22.56 ± 1.93 for SynGAN, 21.74 ± 1.97 for CycleGAN, 23.09 ± 2.26 for U-Net, 22.67 ± 2.28 for WATNet. The Mean SSIM values for the three methods are: 0.7230 ± 0.0496 for SynGAN, 0.6782 ± 0.0413 for CycleGAN, 0.7688 ± 0.0558 for U-Net, 0.7207 ± 0.0490 for WATNet. PSNR = peak signal-to-noise ratio; SSIM = structure similarity.
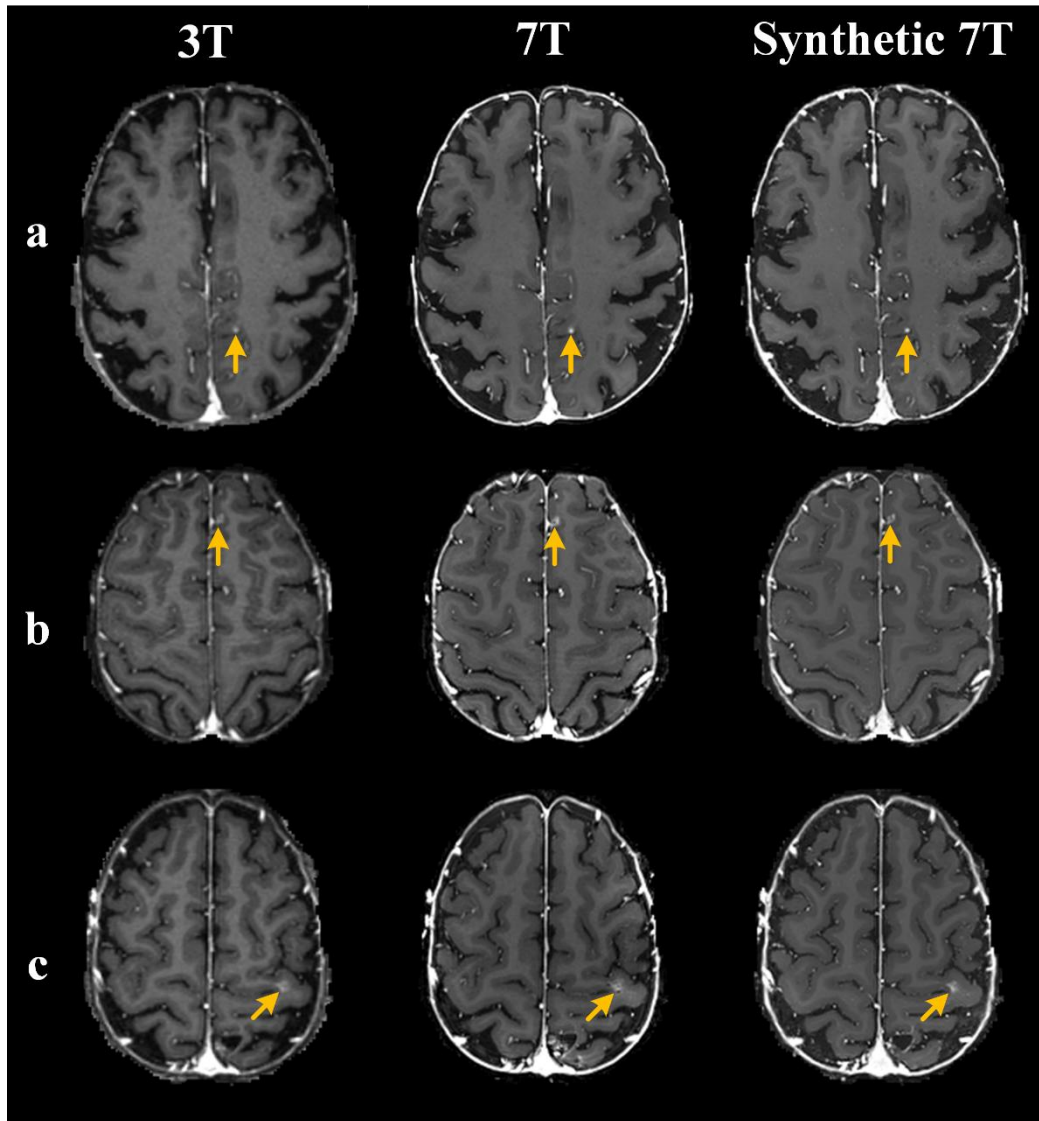
Supplementary Fig. 5: Comparisons of Representative 3T, 7T and Synthetic 7T Images. a, T1-weighed images in a 70-year-old patient with brain metastases from lung cancer. b, T1-weighed images in a 52-year-old patient with brain metastases from lung adenocarcinoma. c, T1-weighed images in a 54-year-old patient with brain metastases from lung adenocarcinoma. The golden arrows indicate the enhanced micrometastases.
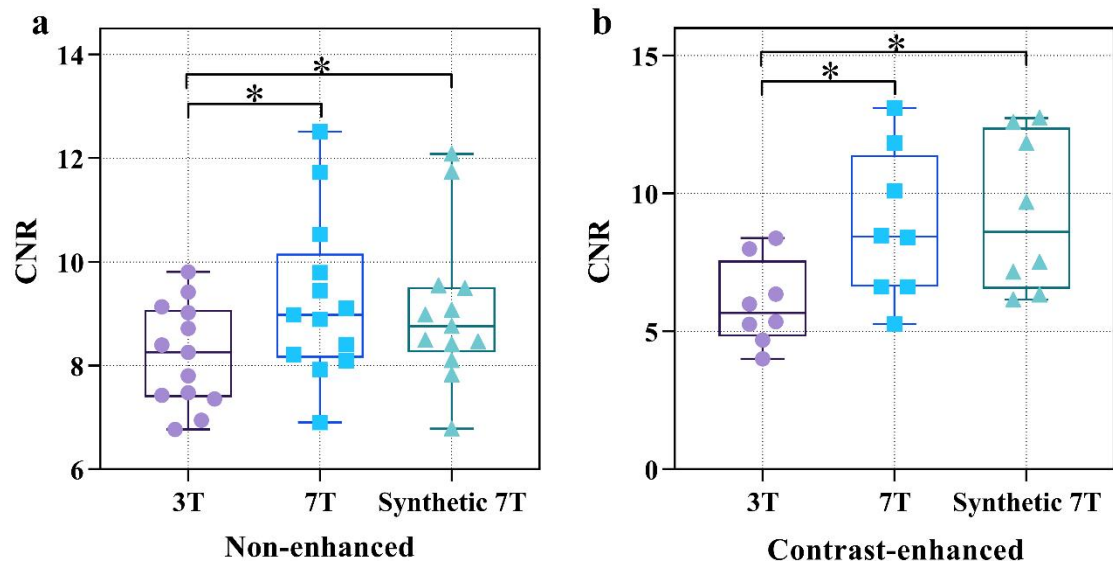
Supplementary Fig. 6: Boxplots of CNR values of the 3T, 7T, and Synthetic 7T images for non-enhanced (a) and contrast-enhanced (b) test data. Each circular dot indicates the CNR value from each subject. The synthetic 7T images showed significantly higher CNR than the 3T images (*P* < 0.05), with no significant difference compared with the real 7T images (*P* > 0.05). *Statistically different results with *P* < 0.05. CNR = contrast-to-noise ratio.