Generative Adversarial Model for MR Image Generation

The model was based on recent works for image-to-image translation using generative adversarial model (*1*). It successfully applied to tasks for synthesizing complex images from simple representations, for instance, translating sketches to photos and colorization of black-and-white images (*1*). We applied this approach to translate PET images to structural MR images. The training process uses generative adversarial network. The generator *G*, a convolutional neural network based on U-net (*2*), is trained to translate PET (*x*) to MR (*y*) images which cannot be distinguished from real images. In contrast, the discriminator *D*, another convolutional neural network, attempts to minimize the misclassification error for distinguishing real pairs of PET/MR images from synthetic pairs of PET/MR images (Fig. 1). Specifically, the objective of the model is

$$L_{Adv}(G,D) = \mathbb{E}_{x,y\ \sim\ p(x,y)}[\log D(x,y)] + \mathbb{E}_{x\ \sim\ p(x)}[1 - \log D(x,G(x))]$$

where $\mathbb{E}_{x,y\ \sim\ p(x,y)}$ represents the expectation that PET (*x*) and MR (*y*) images are sampled from the probability distribution of real pairs *p(x,y)*. $\mathbb{E}_{x,y\ \sim\ p(x,y)}[\log D(x,y)]$ is maximized when $D(x,y)$ = 1 as the output of *D* has the range of 0 to 1. $\mathbb{E}_{x\ \sim\ p(x)}$ represents the expectation that PET is sampled from the probability distribution of *p(x)*. The posterior part is maximized when $D(x,G(x))$=0, while it is minimized when *G* successfully fools the *D*, $D(x,G(x))$=1. Thus, the training of *D* is aimed to maximize the L_{GAN}, while *G* attempts to minimize the L_{GAN}. The adversarial objective function was additionally combined with the loss for pixel-based regression represented by L1-distance between real and generated images.

$$L_{L1}(G,D) = \mathbb{E}_{x,y\ \sim\ p(x,y)}(\| y - G(x) \|_1)$$

Thus, the final objective function is $L(G,D) = L_{Adv} + \alpha L_{L1}$. α is a factor which determines the contribution of two types of losses. In our model, $\alpha = 100$ was used.

Training and Testing for MR Image Generation

Neural networks consisted of generator and discriminator. Specific parameters of the neural networks are described in Supplementary Tables. The generator consisted of multiple convolutional layers with skipped connections, which was based on U-net (*2*). Axial slices of PET images with 192 x 256 matrix size were subsampled to half-sized feature maps iteratively. These convolutional layers finally produced feature maps with 3 x 4 matrices followed by iterative deconvolutional layers which finally produce original sized matrices. Each input of deconvolutional layer was connected to same sized feature matrices produced by a previous convolutional layer. Discriminator consisted of 6 convolutional layers and finally produced 6 x 8 matrix to perform patch-based learning.

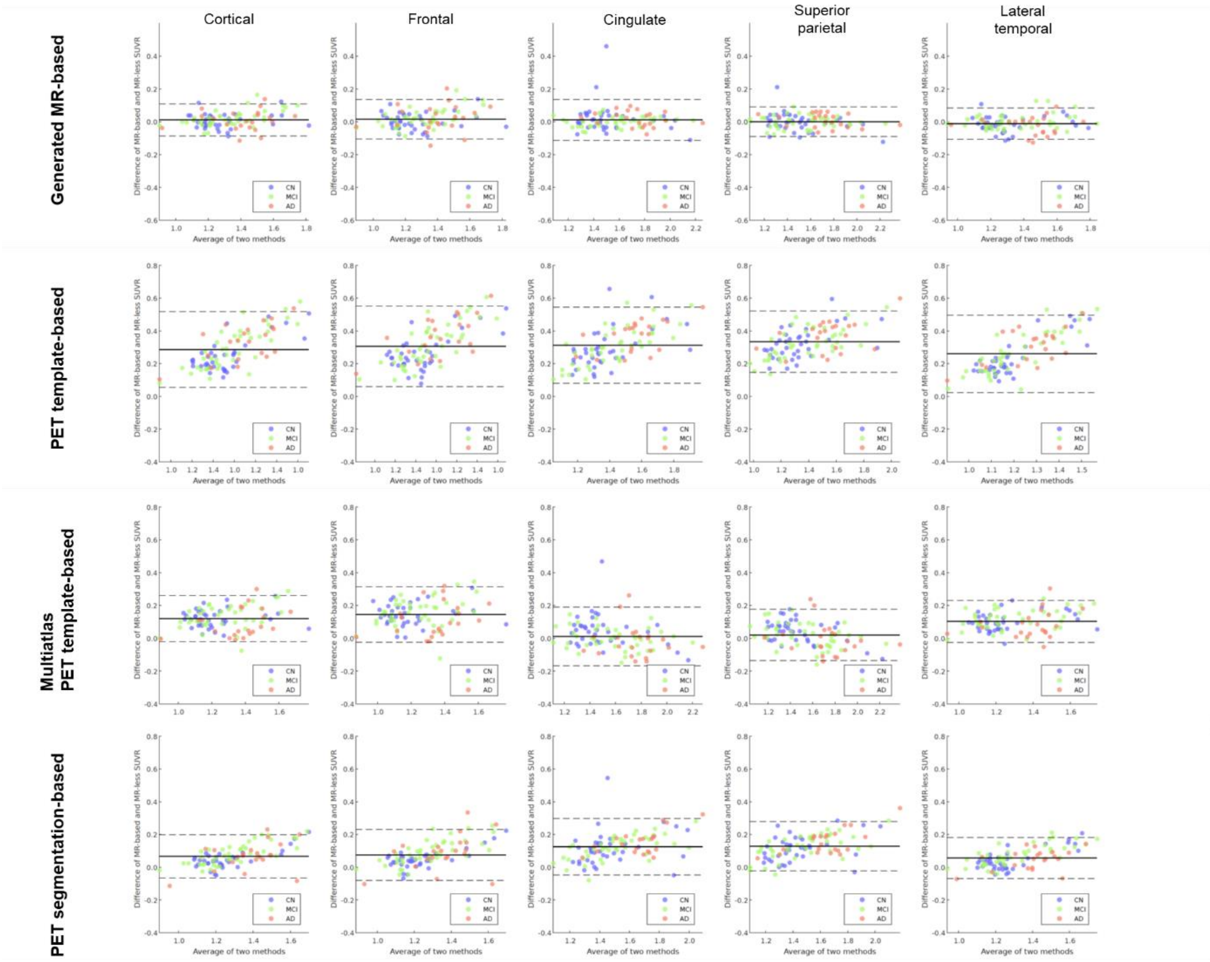
Training was performed by PET and MR images of 169 subjects. Input of the model was an axial slice image of PET scan. Total axial slices of 169 subjects were 32659 images. The network was trained using the Adam optimizer (*3*). Testing was performed by images of 98 subjects independent from the training data. Row PET images were resliced to have voxel size of $1.2 \times 1.0 \times 1.0\ mm^3$ as aforementioned. Generator produced synthetic MR slices corresponding to PET slices.

Supplemental References

1. Isola P, Zhu J-Y, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. *arXiv preprint arXiv:161107004.* 2016.

2. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Paper presented at: International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015.

3. Kingma D, Ba J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:14126980.* 2014.



Supplemental Fig. 1. Bland-Altman plots for different MR-less methods. Bland-Altman plots were drawn for the comparison between the MR-less and MR-based quantification methods. SUVRs calculated using the generated MR-based method showed the most accurate and precise quantification results compared with the other three PET template-based methods. Of note, the normal PET template-based and PET segmentation-based methods showed a trend of higher bias when amyloid load was increased.

Supplemental Fig. 2. Artifacts of generated MR. Generated MR showed artifacts which could not be found in true MR images. In particular, irregular and inhomogeneous signal intensities were found in several generated MR images.

Supplemental Table 1. Network architectures for the generative network, U-net.

	Filter size	Filter number	Activation	Subsampling	Output Size	
Convolutional layer 1	3x3	64		1/2	96x128	Input: Axial slices of PET image
Convolutional layer 2	3x3	128		1/2	48x64	
Convolutional layer 3	3x3	256		1/2	24x32	
Convolutional layer 4	3x3	512		1/2	12x16	
Convolutional layer 5	3x3	512		1/2	6x8	
Convolutional layer 6	3x3	512	LeakyReLU activation	1/2	3x4	
Deconvolution layer 1	3x3	512		2	6x8	Concatenation to Convolutional layer 5
Deconvolution layer 2	3x3	512		2	12x16	Concatenation to Convolutional layer 4
Deconvolution layer 3	3x3	256		2	24x32	Concatenation to Convolutional layer 3
Deconvolution layer 4	3x3	128		2	48x64	Concatenation to Convolutional layer 2
Deconvolution layer 5	3x3	64		2	96x128	Concatenation to Convolutional layer 1
Deconvolution layer 6	3x3	1	Tanh activation	2	192x256	Output: Axial slices of PET image

Supplemental Table 2. Network architectures for the discriminative network.

	Filter size	Filter number	Activation	Subsampling	Output Size	
Convolutional layer 1	5x5	64		1/2	96x128	Input: Pairs of PET and MR images
Convolutional layer 2	5x5	64		1/2	48x64	
Convolutional layer 3	5x5	128	LeakyReLU activation	1/2	24x32	
Convolutional layer 4	5x5	256		1/2	12x16	
Convolutional layer 5	5x5	512		1/2	6x8	
Convolutional layer 6	5x5	1	Sigmoid activation	-	6x8	Output: 1 for real MR, 0 for generated MR