NeuroImage 232 (2021) 117890

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Translating amyloid PET of different radiotracers by a deep generative model for interchangeability ☆

Seung Kwan Kang [a,b], Hongyoon Choi [a,c,*], Jae Sung Lee [a,b,*], for the Alzheimer's Disease Neuroimaging Initiative

[a] Department of Nuclear Medicine, Seoul National University Hospital, 101 Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea
[b] Department of Biomedical Sciences, Seoul National University College of Medicine, 101 Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea
[c] Department of Nuclear Medicine, Seoul National University College of Medicine, 101 Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea

ARTICLE INFO

Keywords:
Amyloid
Alzheimer's disease
Deep learning
Generative model
PET

ABSTRACT

It is challenging to compare amyloid PET images obtained with different radiotracers. Here, we introduce a new approach to improve the interchangeability of amyloid PET acquired with different radiotracers through image-level translation. Deep generative networks were developed using unpaired PET datasets, consisting of 203 [^{11}C]PIB and 850 [^{18}F]florbetapir brain PET images. Using 15 paired PET datasets, the standardized uptake value ratio (SUVR) values obtained from pseudo-PIB or pseudo-florbetapir PET images translated using the generative networks was compared to those obtained from the original images. The generated amyloid PET images showed similar distribution patterns with original amyloid PET of different radiotracers. The SUVR obtained from the original [^{18}F]florbetapir PET was lower than those obtained from the original [^{11}C]PIB PET. The translated amyloid PET images reduced the difference in SUVR. The SUVR obtained from the pseudo-PIB PET images generated from [^{18}F]florbetapir PET showed a good agreement with those of the original PIB PET (ICC = 0.87 for global SUVR). The SUVR obtained from the pseudo-florbetapir PET also showed a good agreement with those of the original [^{18}F]florbetapir PET (ICC = 0.85 for global SUVR). The ICC values between the original and generated PET images were higher than those between original [^{11}C]PIB and [^{18}F]florbetapir images (ICC = 0.65 for global SUVR). Our approach provides the image-level translation of amyloid PET images obtained using different radiotracers. It may facilitate the clinical studies designed with variable amyloid PET images due to long-term clinical follow-up as well as multicenter trials by enabling the translation of different types of amyloid PET.

1. Introduction

Amyloid positron emission tomography (PET) has been widely used to noninvasively evaluate cortical amyloid deposits. Quantitative assessment of amyloid PET is crucial in Alzheimer's disease (AD) given that the amyloid PET is an objective biomarker that predicts the future outcome of patients with mild cognitive impairment (MCI) (Herholz and Ebmeier, 2011). Amyloid PET imaging was initially conducted using [^{11}C]Pittsburgh Compound-B (PIB) (Klunk et al., 2004; Navitsky et al., 2018). However, various ^{18}F-labeled radiotracers, such as florbetapir, florbetaben, and flutemetamol, are widely used in the clinic now (Clark et al., 2011; Curtis et al., 2015; Sabri et al., 2015). These recently introduced ^{18}F-labeled radiotracers have shown comparable performance to [^{11}C]PIB in estimating cortical amyloid load with better availability in terms of half-life of the radioisotope. However, the limited interchangeability between the radiotracers caused by the different kinetics, acquisition setting, and analysis methods prevents long-term clinical trials as well as multicenter comparisons (Villemagne et al., 2012; Wolk et al., 2012). For example, as cognitive disorders need to be managed with long-term follow-up, many patients previously recruited for clinical trials before the wide use of ^{18}F-labeled radiotracers consequently only have [^{11}C]PIB PET data. In these patients, it is difficult to evaluate the longitudinal changes in amyloid deposits due to the lack of interchangeability of the radiotracers.

To overcome the limited interchangeability, recent studies have investigated the relationship between the quantitative measures on PET scans with [^{11}C]PIB and ^{18}F-labeled radiotracers (Landau et al., 2013;

☆ Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

* Corresponding author.
E-mail addresses: chy1000@snu.ac.kr (H. Choi), jaes@snu.ac.kr (J.S. Lee).

https://doi.org/10.1016/j.neuroimage.2021.117890
Received 22 June 2020; Received in revised form 31 December 2020; Accepted 15 February 2021
Available online 19 February 2021

Table 1
Demographics and clinical diagnosis of training and test dataset.

	[11C]PIB PET training dataset (n = 203)	[18F]florbetapir PET training dataset (n = 850)	Paired test dataset (n = 15)
Age	73.2 ± 7.5 (55.0 - 87.0)	72.3 ± 7.2 (55.0 - 91.0)	73.8 ± 5.2 (62.0 – 81.0)
Sex (M:F)	133:70	443:407	11:4
Diagnosis (AD:MCI:NC)	48:115:40	250:76:172	5:8:2

Villemagne et al., 2012). In addition, a working group has proposed a 'Centiloid' scaling method to standardize the quantitative amyloid deposit estimation by PET. In the Centiloid scaling method, typical amyloid-negative PET scans have a mean of zero Centiloid unit and typical amyloid-positive PET scans score on average 100 regardless of the radiotracers used (Klunk et al., 2015). This method is relatively easy to use for scaling amyloid PET images acquired using the ^{18}F-labeled radiotracers. However, voxel-wise translation at the image level is limited because the linear Centiloid scaling parameters for all the voxels are the same.

To date, the direct translation of different amyloid PET images to achieve better interchangeability has not been attempted yet. Rapidly advancing deep generative models allow the synthesis of new images from a given image in various imaging applications (Goodfellow et al., 2014; Zhu et al., 2017). Recently, the image generation based on the deep neural networks has been used for amyloid PET quantification without the help of structural MR images and PET attenuation correction without CT-based attenuation maps (Choi and Lee, 2018; Hwang et al., 2018; Kang et al., 2018; Lee, 2020; Leynes et al., 2017). However, the generation of different amyloid PET images is challenging because of the limited number of patients who underwent paired amyloid PET scans (e.g. [11C]PIB and [18F]florbetapir) performed in short interval. Here, we propose a method to translate between [11C]PIB and [18F]florbetapir PET images by applying deep generative networks to achieve the radiotracer interchangeability at the image level.

2. Materials and methods

2.1. Patient population

In this study, image data were collected from the Alzheimer's Disease Neuroimaging Initiative (ADNI) (http://adni.loni.usc.edu) database. The ADNI was launched in 2003 as a public-private partnership led by principal investigator Michael W. Weiner, MD, VA Medical Center and University of California, San Francisco. The ADNI includes subjects from over 50 sites across the US and Canada. The primary goal of the ADNI is to develop combined biomarkers by testing whether serial imaging, biological markers and clinical, and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment and early Alzheimer's disease. For up-to-date information, see http://www.adni-info.org.

Eight hundred fifty subjects who underwent [18F]florbetapir PET as a baseline amyloid imaging study were included as a training dataset. Another dataset consisting of 203 [11C]PIB PET scans were also downloaded from the ADNI. [11C]PIB PET scans obtained both at baseline and at follow-up visits were included in this cohort. Notably, we trained a deep generative model on all [18F]florbetapir and [11C]PIB PET scans regardless of the acquisition site. The demographic data of the two datasets are summarized in **TABLE 1**. To validate our model, we used the data from the patients who underwent both [11C]PIB and [18F]florbetapir PET scans. Fifteen patients had both PET scan data obtained with 1- or 2-year intervals. Note that these [11C]PIB and [18F]florbetapir PET scans were not used for the training.

2.2. Image preprocessing

To train and test the networks, minimally preprocessed ADNI PET data were used. Only coregistration and averaging were conducted on the data (second type). These PET data were acquired from different hardware systems and reconstructed with different parameters. Both the [18F]florbetapir and [11C]PIB image data consisted of four 5-min frames acquired 50–70 min after the tracer injection. To minimize image blurring due to head motion, image registration was applied to the 5-min dynamic frames (Jagust et al., 2015). Because no spatial transformation was performed, all PET images had different shapes and variable voxel sizes. The images were resampled to have a voxel size of $2 \times 2 \times 2$ mm^3 using trilinear interpolation. This processing was applied to both training and test data. We started with these images for proposed task and its evaluations. Further image transformation was not performed; thus, all PET images had different shapes and positions. To stabilize the training, we also normalized the intensity of each input image to range from −1 to 1. The images were rescaled back at the testing phase.

2.3. Development of generative model

The cycle generative adversarial network (CycleGAN) structure was employed to translate the unpaired amyloid PET images (Zhu et al., 2017). **FIG. 1** shows a detailed structure of the proposed networks. There are two generators and two discriminators. A generator converted [11C]PIB images to [18F]florbetapir images and another generator converted [18F]florbetapir images to [11C]PIB images. Each discriminator was trained to differentiate real PET images from generated images. For the training of CycleGAN, the following three losses were employed (Zhu et al., 2017). The first loss is the GAN loss retaining the generative characteristics of the neural networks (Goodfellow et al., 2014).

$$\begin{aligned} L_{GAN}(G_1, G_2, D_1, D_2) &= L_{GAN}^{PIB}(G_1, D_1) + L_{GAN}^{AV45}(G_2, D_2) \\ L_{GAN}^{AV45}(G_1, D_1) &= \mathbb{E}_{b\sim p_{AV45}}\left[log D_1(b)\right] + \mathbb{E}_{a\sim p_{PIB}}\left[\log\left(1 - D_1(G_1(a))\right)\right] \\ L_{GAN}^{PIB}(G_2, D_2) &= \mathbb{E}_{a\sim p_{PIB}}\left[log D_2(a)\right] + \mathbb{E}_{b\sim p_{AV45}}\left[\log\left(1 - D_2(G_2(b))\right)\right] \end{aligned} \quad (1)$$

where a and b are the sampled elements from each probabilistic distribution of images, [11C]PIB and [18F]florbetapir, respectively. The G_1 and G_2 are the generators that translate amyloid images, and the D_1 and D_2 are the discriminators in the CycleGAN structure. The second loss is the reconstruction loss comparing the output of the second generator from the input. This loss enables the output of the generative model to be similar to the input of the networks.

$$\begin{aligned} L_{cyc}(G_1, G_2) &= L_{cyc}^{PIB}(G_1) + L_{cyc}^{AV45}(G_2) \\ L_{cyc}^{PIB}(G_1) &= \mathbb{E}_{b\sim p_{AV45}}\left[\left\|b - G_1\left(G_2(b)\right)_1\right\|\right] \\ L_{cyc}^{AV45}(G_2) &= \mathbb{E}_{a\sim p_{PIB}}\left[\left\|a - G_2\left(G_1(a)\right)_1\right\|\right] \end{aligned} \quad (2)$$

The final one was an identical loss. Followed by this loss, each generator should produce similar images to the given input that is from real samples. The authors of (Zhu et al., 2017) showed that the results were more robust with an identical loss.

$$\begin{aligned} L_{id}(G_1, G_2) &= L_{id}^{PIB}(G_1) + L_{id}^{AV45}(G_2) \\ L_{id}^{PIB}(G_1) &= \mathbb{E}_{b\sim p_{AV45}}\left[\|b - G_1(b)_1\|\right] \\ L_{id}^{AV45}(G_2) &= \mathbb{E}_{a\sim p_{PIB}}\left[\|a - G_2(a)_1\|\right] \end{aligned} \quad (3)$$

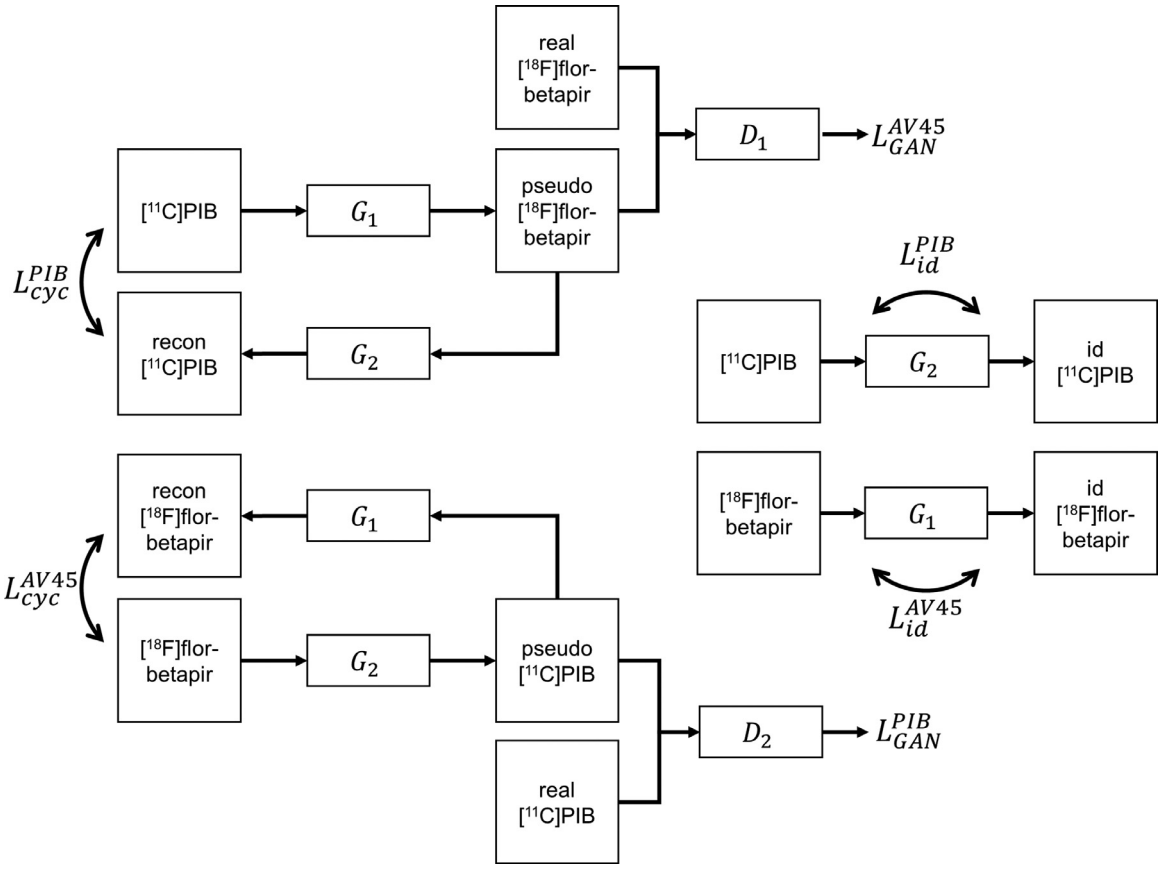

Fig. 1. Schematic of the proposed CycleGAN structure. Three losses indicated in the figure correspond to Eqs. (1), (2), (3) and (4).

The above three losses are combined with the following hyperparameters determined empirically.

$$L = L_{GAN} + 100 \times L_{cyc} + 10 \times L_{id} \tag{4}$$

We combined DenseNet and U-Net architectures and used them as a generator (Huang et al., 2016; Ronneberger et al., 2015). Detailed network structures are shown in **Supplementary Table 1**. We also applied the spectral normalization method to the convolutional operation of the discriminator to stabilize the training of the GAN (Miyato et al., 2018). Three-dimensional image, not two-dimensional slices, was fed into the networks producing the pair of amyloid PET images to consider the regional PET activity distribution. To augment the number of data set for training the network, the input of the network was randomly rotated from −10 to 10° and flipped for each of the three-axes. The input image dimension was $64 \times 64 \times 36$, which was two-fold downsampled from the original image after cropping useless margin to reduce the memory usage of the neural networks. The same number of imaging data randomly selected from each of the datasets, [^{11}C]PIB and [^{18}F]florbetapir, was used for each mini-batch. The number of mini-batch was one, and we trained the network for 1000 epochs. The learning rate used for training the neural network was 0.0001.

2.4. Evaluation of amyloid PET scans

After training proposed neural network, the results were reviewed by a 10-year experienced nuclear medicine physician (H.C). To validate our model for translating amyloid PET images with different radiotracers, global and regional SUVR estimated from the paired samples were compared using cerebellum gray matter as a reference region. We used 15 paired [^{11}C]PIB and [^{18}F]florbetapir PET images as well as the corresponding 3D structural T1 MRI images scanned closest to the [^{11}C]PIB except for one subject (**Supplementary Table 2**). Images were translated into their individual space by the neural network. For the visual comparisons, [^{11}C]PIB images were co-registered to the [^{18}F]Florbetapir images using the SPM12. This process was performed in the subject's individual space for the visualization. For the quantitative analyses, both sets of PET images were spatially normalized using the deformation fields estimated from the same MRI by statistical parametric mapping (SPM12) (Ashburner, 2007; Ashburner and Friston, 2005). The PET registration and spatial normalization of [^{11}C]PIB and [^{18}F]florbetapir data were conducted separately to the MRI. We extracted regional PET counts from five regions, which were the composite VOIs of the global cerebral cortex, frontal cortex, cingulated cortex, superior parietal cortex, and lateral temporal cortex defined in the AAL template (Tzourio-Mazoyer et al., 2002) after spatially normalizing each image to the MNI space. The counts of the five VOIs were divided by the counts of the cerebellum to calculate the SUVR. We also generated subject-space SUVR images using registered [^{11}C]PET and [^{18}F]Florbetepir via dividing registered images by extracted cerebellum counts. Linear regression was performed between the SUVRs obtained from the generated images and those obtained from the original images. Additionally, we calculated the intraclass correlation coefficients (ICCs) of the regional SUVRs between the generated and original images to assess the consistency of the measurements (McGraw and Wong, 1996):

$$ICC = \frac{V_b}{V_a + V_b}, \tag{5}$$

where V_a is the variance between methods and V_b is the variance between measurements. The denominator equals to the total variance of the data.

We also conducted voxel-wise comparisons using paired *t*-test in the SPM12 after smoothing SUVR images with an isotropic 3D Gaussian

filter of an 8 mm full width at half maximum (FWHM). The voxel size was isotropic 2 mm. Thresholds of $p\langle$ 0.0001 and $k\rangle$ 500 contiguous voxels were applied to the results. Multiple comparison correction was not conducted because of the small number of the test dataset.

For the comparison, we conducted the Centiloid scaling of [^{11}C]PIB and [^{18}F]florbetapir PET images. The results of the replication of level-1 analysis in Klunk et al. were given in **Supplementary Figure 1**. Standard Centiloid cortical and whole cerebellar VOI templates were applied to the spatially normalized PET images to obtain SUVR. The SUVR values for each tracer were then transformed into Centiloid units by the following linear equations (Klunk et al., 2015; Kolibash, accessed 12 August 2020):

$$CL_{PIB} = 100 \times \left(SUVR_{PIB} - 1.009\right)/1.068 \quad (6)$$

$$CL_{AV45} = \left(196.9 \times SUVR_{AV45}\right) - 196.03 \quad (7)$$

Additionally, CL scales for non-standard VOIs (above 5 regions) were calculated by multiplying and adding the scaling parameters to them. There were no processing differences among observed PET data, generated pseudo PET images and images used to derive Centiloids.

The visual review of paired PET data was to find whether the generated images changed the visual reading in terms of the amyloid positivity. Images were interpreted as binary categories, amyloid positive and negative, without clinical information as well as whether the images were generated or real images.

3. Results

Amyloid PET images corresponding to the administration of different radiotracers were generated from original PET images. Pseudo-[^{18}F]florbetapir PET images were generated from [^{11}C]PIB PET images, and pseudo-[^{11}C]PIB PET images were generated from [^{18}F]florbetapir PET images using the neural networks. **FIG. 2** shows the representative pseudo-[^{18}F]florbetapir and pseudo-[^{11}C]PIB PET images for an amyloid-negative case. It also shows the generated PET images for an amyloid-positive case. The original [^{11}C]PIB uptake pattern for the amyloid-positive case is largely different from the original [^{18}F]florbetapir uptake pattern due to differences in the binding affinities. Nonetheless, the generated PET images reflect these different patterns.

To validate the generated PET images and assess the radiotracer interchangeability, we performed a head-to-head comparison between the original and generated PET images in terms of the SUVRs (**FIG. 3**). Regional and global SUVRs were estimated from five predefined VOIs (the global cerebral cortex and the frontal, cingulate, superior parietal, and lateral temporal cortices). As shown in **FIG. 3 A**, the SUVR calculated from the original [^{18}F]florbetapir is relatively lower than that calculated from the original [^{11}C]PIB, which is consistent with the results from a previous report (Klunk et al., 2015). The calculated ICC between original and pseudo-[^{18}F]florbetapir images for the global cortex was 0.87. Likewise, the ICC between original and pseudo-[^{11}C]PIB images for the global cortex was 0.85; however, the ICC result between the original [^{11}C]PIB and [^{18}F]florbetapir images yielded a lower value, 0.65 (**TABLE 2**). In the Centiloid cortical VOI, our method yielded comparable ICC results to the Centiloid-scaled data (0.94 for original [^{18}F]florbetapir vs. pseudo-[^{18}F]florbetapir, 0.91 for original [^{11}C]PIB vs. pseudo-[^{11}C]PIB, and 0.93 for Centiloid-scaled [^{18}F]florbetapir vs. Cetiloid-scaled [^{11}C]PIB). The ICC between the original image pairs, [^{18}F]florbetapir and [^{11}C]PIB PET, was only 0.45 in the Centiloid VOIs. Moreover, the ICC values on deep gray showed that the Centiloid scaling did not improve the correlation. On the other hand, our method improved the correlation in the deep gray as a region with relatively low amyloid deposits (**Supplementary Table 3**). The five ROIs except for Centiloid VOI were based on the cerebellum gray reference region and the results using Centiloid reference region (whole cerebellum) were also given in **Supplementary Table 4**. The slope from the linear regression performed between original and pseudo-[^{18}F]florbetapir PET was 1.092 for the global cortex SUVR although some samples deviated from the perfect agreement. The slope from the linear regression performed between the original and pseudo-PIB PET data was 1.043. The SUVR subtraction images between [^{11}C]PIB and [^{18}F]florbetapir PET and their corresponding generated images are presented in **FIG. 4**. The relatively lower SUVR in the original [^{18}F]florbetapir PET than the original [^{11}C]PIB PET are noticeable in the amyloid-positive case, which is in line with **FIG. 3 A**, and this difference was remarkably reduced in the generated PET images. Bland-Altman plots for the SUVRs of the paired PET data are shown in (**FIG. 5**). The Bland-Altman plot for Centiloid scaled data for standard Centiloid VOI was also presented in **Supplementary Figure 2**. The comparison between the original [^{11}C]PIB and [^{18}F]florbetapir PET images showed relatively high bias and high variance. Quantification using the generated PET data, the pseudo-[^{18}F]florbetapir and pseudo-[^{11}C]PIB PET data, showed almost unbiased results for all types of VOIs. These findings were confirmed in the voxel-wise statistical analysis shown in **FIG. 6**.

4. Discussion

In this study, we proposed a deep generative model trained using unpaired data to translate amyloid PET images acquired with different radiotracers. The emergence of deep learning with convolutional neural networks has led to many innovations in image processing fields as well as in medical imaging (Choi and Lee, 2018; Hegazy et al., 2019; Hwang et al., 2018; Lee, 2020; Nie et al., 2016; Park et al., 2018; Prasoon et al., 2013; Ronneberger et al., 2015; Yi et al., 2019; Yonekura et al., 2018). The deep learning applications were mainly made through supervised learning methods requiring paired datasets between inputs and labels, which are difficult to collect in a real clinical environment. Accordingly, generative modeling that does not require explicit labels has received much attention (Yi et al., 2019). In particular, image generation using unpaired data allows the utilization of various real medical image data given that the unpaired data are widely available in the hospitals.

Our suggested method could provide an image-level translation that enables a comparison of amyloid PET images acquired with different radiotracers. Previously, [^{11}C]PIB PET had been used before the wide use of amyloid PET scans with [^{18}F]-labeled radiotracers. As cognitive disorders require long-term follow-up, particularly for clinical trials, an interchangeable quantification of amyloid deposits using different radiotracers is required when patients have had previous [^{11}C]PIB PET images. Although the recent Centiloid scale provides an interchangeable and common quantification scale, it is limited to linear scaling of all voxels of SUVR images. Amyloid deposits in the brain have a specific spatial pattern that is associated with disease progression (Hanseeuw et al., 2018). Moreover, subsets of dementia show different distribution patterns of amyloid deposits (Frings et al., 2015; Mesulam et al., 2014). Therefore, there are several advantages of regional SUVR calculation and image voxel-level analyses over the normalized global SUVR as image-based biomarkers. Our results showed that the global SUVR evaluation is also possible using the deep generative model with a similar performance to the Centiloid scaling. Furthermore, our method allows the application of other atlases and VOIs for the regional analysis (**TABLE 2**). A few cases in this study showed minor differences in the detailed radiotracer distribution patterns between [^{11}C]PIB PET and [^{18}F]florbetapir PET. For instance, the trained network detected regional uptake differences in areas such as the occipital lobes (**FIG. 2**, indicated with yellow arrows). In this regard, the quantification of the regional amyloid deposit pattern, as well as the spatial differences rather than the simple global SUVR estimation, would be important in the assessment of amyloid PET. Besides, the ICC values for deep gray matter regions shows how the two methods, Centiloid and the proposed method, could result

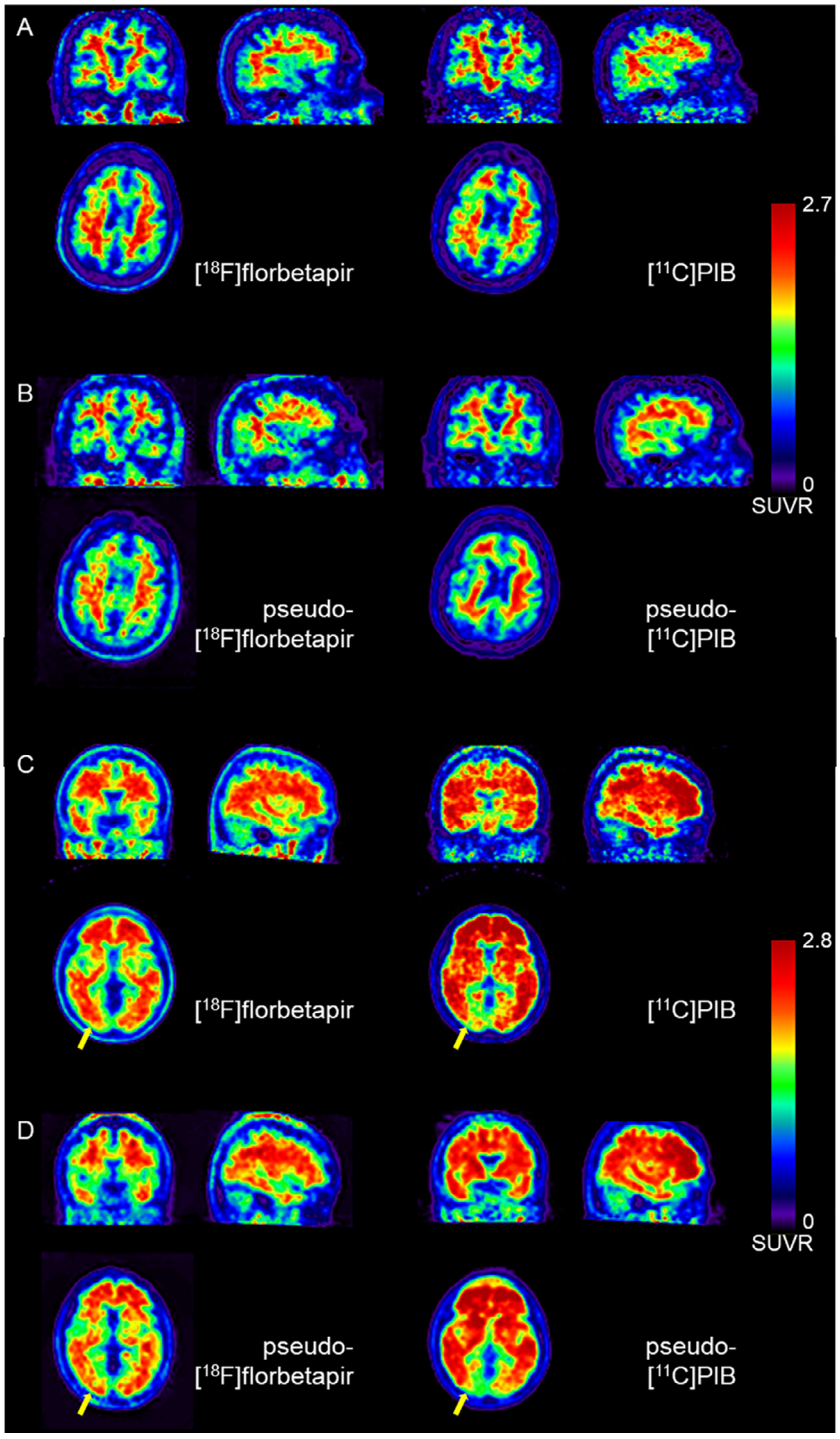

Fig. 2. Qualitative results from paired subjects. (A) The original [^{18}F]florbetapir and [^{11}C]PIB images for an amyloid-negative case, (B) a pseudo-[^{18}F]florbetapir image generated from the [^{11}C]PIB PET images and a pseudo-[^{11}C]PIB image generated from the [^{18}F]florbetapir PET images for the amyloid-negative case, (C) the original [^{18}F]florbetapir and [^{11}C]PIB images for an amyloid-positive case and (D) a pseudo-[^{18}F]florbetapir image generated from the [^{11}C]PIB PET images and a pseudo-[^{11}C]PIB image generated from the [^{18}F]florbetapir PET images for the amyloid-positive case. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

in different quantification results in brain regions with relatively low specific bindings (**Supplementary Table 3**) where the amyloid deposition in these regions can be used as a diagnostic measure of disease progression (Cho et al., 2018; Hanseeuw et al., 2018).

The translation of amyloid PET with different radiotracers could facilitate multicenter and long-term follow-up clinical studies by enabling the use of various amyloid PET images. Our approach could provide information on the temporal changes in the amyloid distribution pattern in the brain for each subject if the subject undergoes longitudinal imaging with amyloid PET using [^{18}F]-labeled radiotracers as well as [^{11}C]PIB. Recently, various [^{18}F]-labeled radiotracers have been used for the assessment of amyloid deposits in many centers. The same method proposed in this study can be applied to the translation between amyloid PET scans with different [^{18}F]-labeled radiotracers. As amyloid PET has become a more widely available test, studies based on retrospective, multicenter, and large PET data with various tracers may be feasible when our model is applied.

In this study, the model responsible for the translation of amyloid PET was trained and validated using PET data from multiple centers of the ADNI, which include different PET protocols and machines. We trained the model using PET images with minimally processed data to generalize the model. Various preprocessing steps, including image registration and Gaussian smoothing for image harmonization (Jagust et al., 2015), are available in the ADNI. However, to develop a generalized

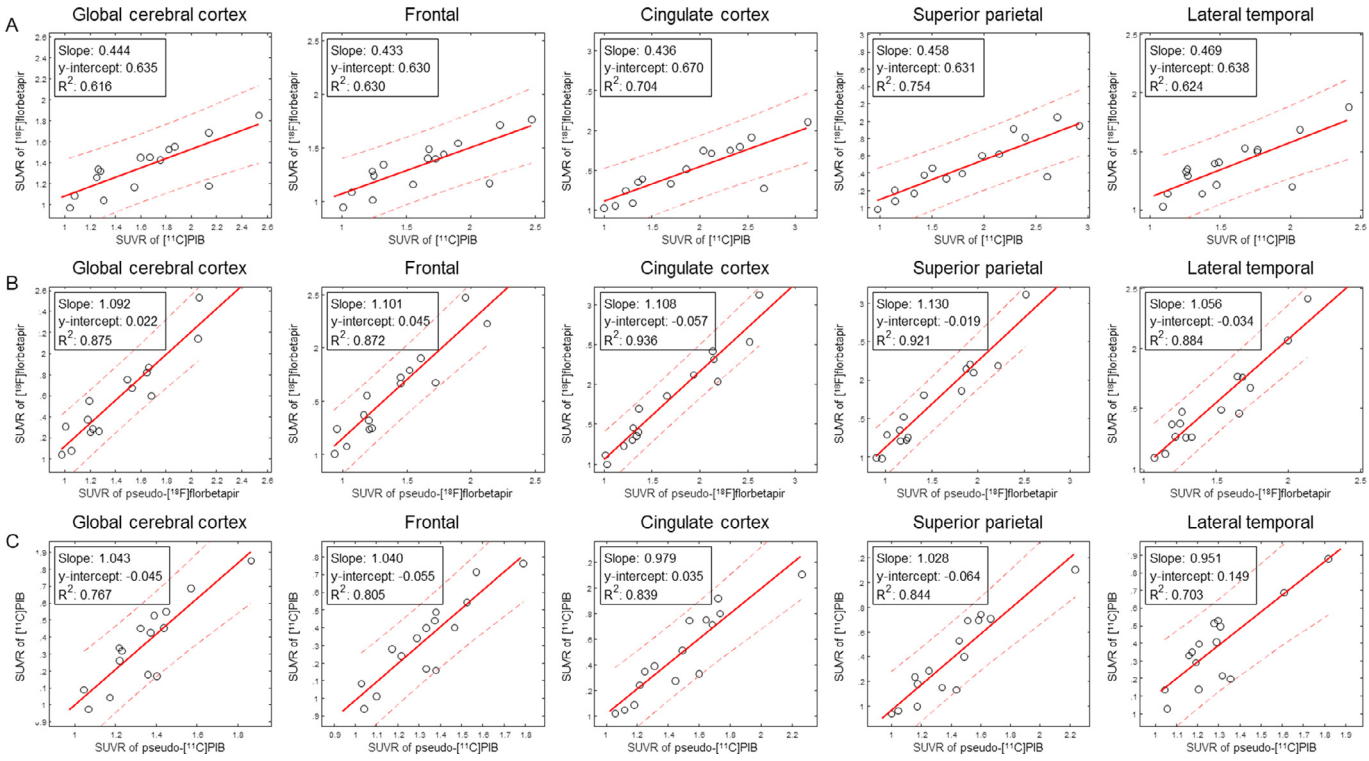

Fig. 3. Correlation plots for the regional SUVRs between paired [^{11}C]PIB and [^{18}F]florbetapir data and between the original data and the generated images. Each column shows different regions: the global cerebral cortex, frontal cortex, cingulated cortex, superior parietal cortex, and lateral temporal cortex. The thick red line is the regression line and dashed red lines are the 95% confidence intervals (CI). (A) Original [^{11}C]PIB vs. original [^{18}F]florbetapir, (B) original [^{18}F]florbetapir vs. pseudo-[^{18}F]florbetapir and (C) original [^{11}C]PIB vs. pseudo-[^{11}C]PIB. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Intraclass correlation calculation for five predefined regions based on AAL template and a Centiloid VOI on the test dataset. The numbers in parentheses mean lower/upper bounds of 95% confidence interval (CI).

	Centiloid VOI	Global cerebral cortex	Frontal	Cingulate cortex	Superior parietal	Lateral temporal
Pseudo-[^{18}F]florbetapir vs. Original [^{18}F]florbetapir	0.94 (0.84 – 0.98)	0.87 (0.66 – 0.95)	0.89 (0.72 – 0.96)	0.92 (0.78 – 0.97)	0.92 (0.77 – 0.97)	0.77 (0.45 – 0.92)
Pseudo [^{11}C]PIB vs. Original [^{11}C]PIB	0.91 (0.76 – 0.97)	0.85 (0.63 – 0.95)	0.82 (0.56 – 0.94)	0.94 (0.83 – 0.98)	0.90 (0.74 – 0.97)	0.93 (0.81 – 0.98)
Original [^{11}C]PIB vs. Original [^{18}F]florbetapir	0.45 (−0.04 – 0.77)	0.65 (0.25 – 0.87)	0.62 (0.18 – 0.85)	0.66 (0.26 – 0.87)	0.50 (0.02 – 0.80)	0.73 (0.39 – 0.90)
Centiloid scaled [^{11}C]PIB vs. Centiloid scaled [^{18}F]florbetapir	0.93 (0.80 – 0.97)	0.82 (0.57 – 0.94)	0.87 (0.66 – 0.95).	0.89 (0.73 – 0.96)	0.88 (0.70 – 0.96)	0.72 (0.35 – 0.89)

model that could be applied to PET images with different protocols and machines, we directly used PET data without harmonization. Our approach of training without data harmonization for the ADNI indicates that the model can be used for PET images obtained from various centers with different hardware and reconstruction algorithms. Nonetheless, since we only applied our model to ADNI data to develop the translation between [^{18}F]florbetapir and [^{11}C]PIB PET, further validation with different tracers is needed. Furthermore, the translation between different types of amyloid PET with ^{18}F-labeled radiotracers is also clinically needed to compare PET scans acquired in multiple centers. As a proof-of-concept study of the image-level translation of amyloid PET, various types of translation can be developed and optimized using unpaired data sets. Note that paired data from two different amyloid PET scans obtained from the same patients are not required for developing translational models based on our approach. Therefore, our generative model can be flexibly customized for the clinical study.

One of the limitations of the study is the training data imbalance: the numbers of [^{18}F]florbetapir and [^{11}C]PIB scans were quite different. This is a frequently occurred problem especially for medical imaging. To resolve the issue, we applied on-the-fly random augmentation in which randomly rotated or flipped images were provided to the networks as input each time, significantly improving the performance. The data imbalance was the intrinsic problem of the ADNI database, nonetheless, the proposed network worked well for this unbalanced dataset. It is because CycleGAN's feature and the advantage are that image-level generation is possible even if the distribution of two datasets is different. Another issue in the training dataset is that [^{11}C]PIB set included longitudinal data for the same participant. However, the follow-up images would not the same data due to the differences in amyloid deposit level

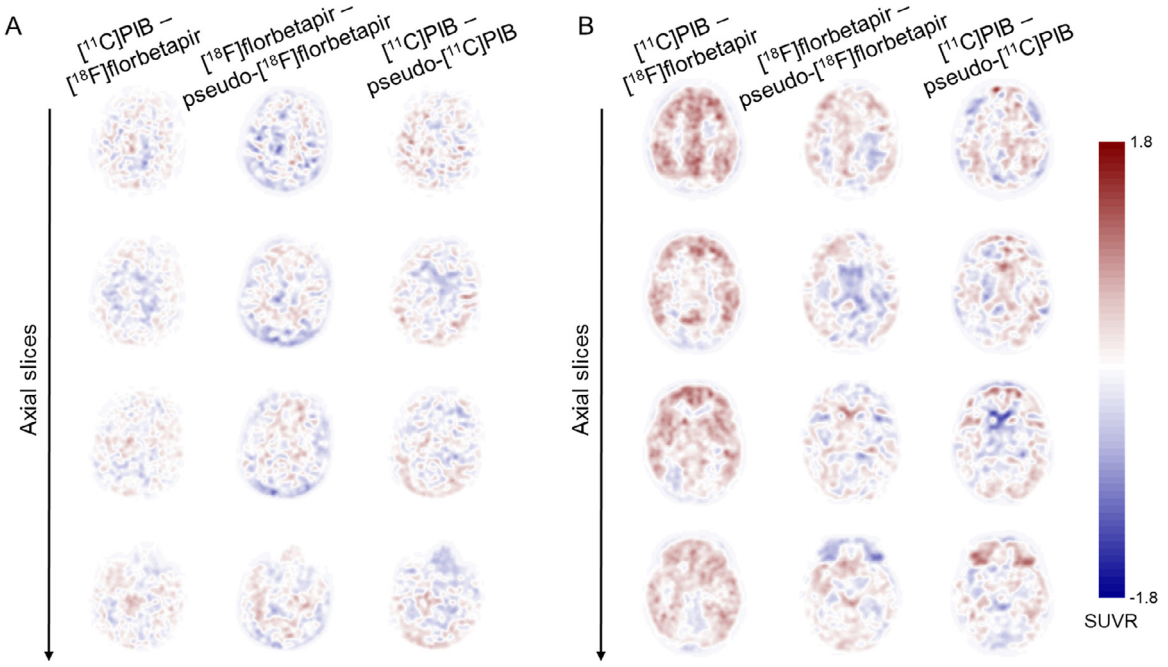

Fig. 4. Difference map for SUVR between the generated image from trained neural networks and labels. For a better visibility, skull stripping was applied to both images. (A) Amyloid negative case and (B) amyloid positive case.

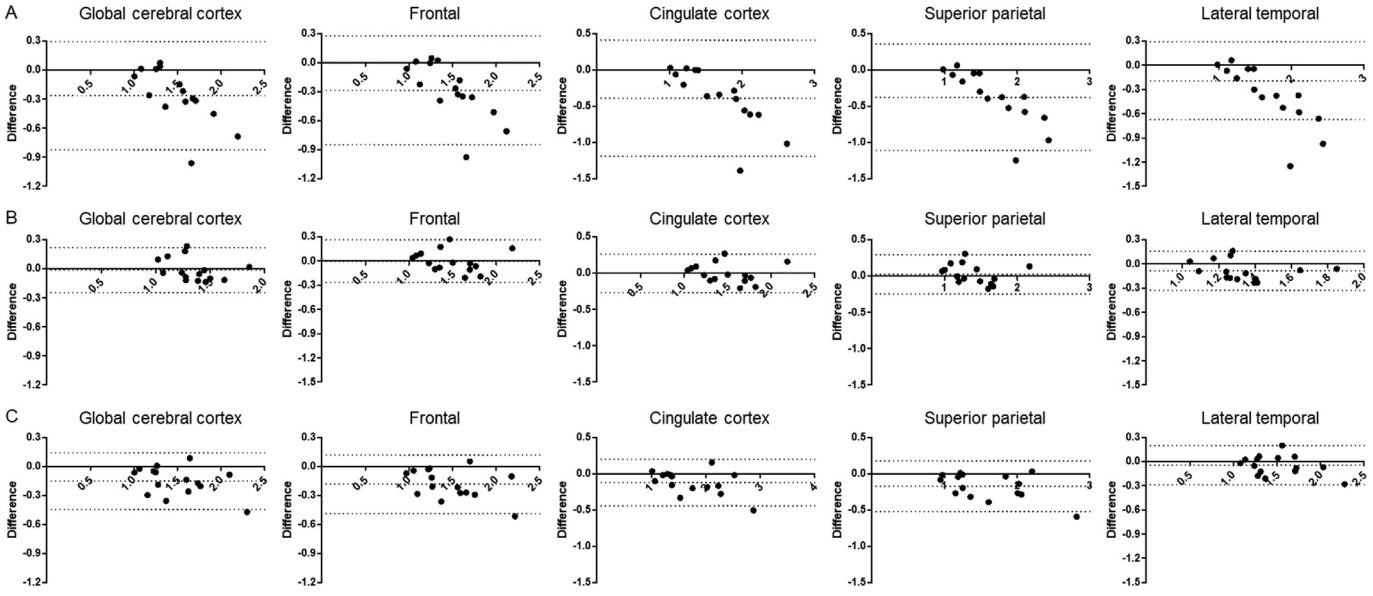

Fig. 5. The Bland plot of regional SUVRs. SUVRs obtained from paired [^{11}C]PIB and [^{18}F]florbetapir and between the original data and the generated data were compared. Each column shows different regions: the global cortex, frontal cortex, cingulated cortex, superior parietal cortex and lateral temporal cortex. The *x*-axis is the average between samples. Dashed lines represent the mean and the range of the 95% confidence intervals of the difference among SUVRs. (A) [^{11}C]PIB vs. [^{18}F]florbetapir, (B) [^{18}F]florbetapir vs. pseudo-[^{18}F]florbetapir and (C) [^{11}C]PIB vs. pseudo-[^{11}C]PIB PET. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and geometrical position and the test-retest variability of the PET scan. Regarding the scan date difference, [^{18}F]florbetapir PET was acquired after [^{11}C]PIB PET because of the protocol of the ADNI cohort (mean: 544.13 days, **Supplementary Table 2**). Therefore, it could be a cause of bias in the quantification of SUVR due to increased amyloid deposits during follow-up.

The other limitation in the application of the generative model to medical imaging is that the generated image is a virtual image that potentially contains artifacts. It is regarded as a hallucinatory effect that creates unwanted structures (Cohen et al., 2018). In this study, we employed the generative model to synthesize pseudo-amyloid PET images acquired with different tracers to achieve the interchangeability in quantification rather than for direct diagnostic purposes. This means that even if there are some structural abnormalities caused by the generative model, large complications do not arise because the major output of the generative model is a quantitative result calculated from predefined VOIs or voxel-wise spatial patterns. We showed that the mean SUVRs obtained from the generated model are virtually equivalent to those obtained from amyloid PET scans acquired with different radiotracers in the independent dataset. Moreover, differently classified amyloid posi-

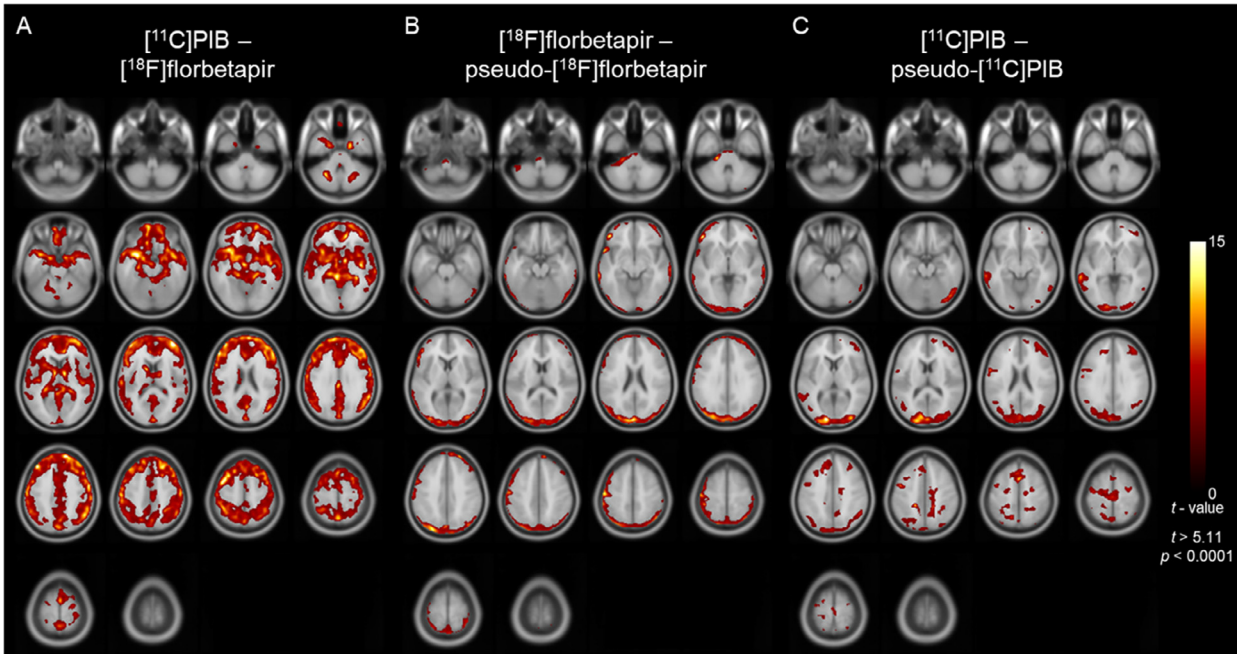

Fig. 6. Voxel-wise comparisons of SUVR images using raw images and pseudo images. (A) [^{11}C]PIB vs. [^{18}F]florbetapir, (B) [^{18}F]florbetapir vs. pseudo-[^{18}F]florbetapir and (C) [^{11}C]PIB vs. pseudo-[^{11}C]PIB PET. ($p\langle$ 0.0001 and $k\rangle$ 500).

tivity or negativity from original images was not observed in the careful visual inspection of all generated images. Thus, at least, the quantification of regional SUVRs could be used for the interchangeable comparison. The usefulness of unpaired datasets provides many opportunities for image translation problems. In future work, the proposed method can be expanded to multi-to-multi-tracer cases using the methods described in (Almahairi et al., 2018; Choi et al., 2018) if the PET images for each dataset are prepared in advance.

Conclusions

In this study, we proposed a deep learning model based on a generative adversarial network to translate amyloid PET images obtained with different radiotracers. The different uptake patterns resulting from the use of the different radiotracers were well generated. The translated PET images could be used for the interchangeable quantification with minimal bias and error. Furthermore, as the model provides image-level estimation, regional SUVR patterns can be assessed in the generated images and compared for different types of amyloid PET images. As a proof-of-concept study, our approach may be extended to various types of amyloid PET images, facilitating clinical studies that require quantification and image-level analyses of different amyloid PET images.

Data availability statement

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.

Author contibutions

H.C. and J.S.L. designed the study. S.K.K. and H.C. developed the model. S.K.K. performed data analyses. All authors contributed to the interpretation of the data and wrote the paper.

Declaration of Competing Interest

The authors declare no potential conflict of interest.

Acknowledgement

This research was supported by the National Research Foundation of Korea (NRF-2019R1F1A1061412 and NRF-2019K1A3A1A14065446). This work was supported by the Korea Medical Device Development Fund grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health & Welfare, the Ministry of Food and Drug Safety) (Project Number: 202011A06) and Seoul R&BD Program (BT200151). Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.;Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.117890.

References

Almahairi, A., Rajeswar, S., Sordoni, A., Bachman, P., Courville, A., 2018. Augmented cyclegan: learning many-to-many mappings from unpaired data. arXiv preprint-arXiv:1802.10151.

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. Neuroimage 38, 95–113.

Ashburner, J., Friston, K.J., 2005. Unified segmentation. Neuroimage 26, 839–851.

Cho, S.H., Shin, J.-.H., Jang, H., Park, S., Kim, H.J., Kim, S.E., Kim, S.J., Kim, Y., Lee, J.S., Na, D.L., Lockhart, S.N., Rabinovici, G.D., Seong, J.-.K., Seo, S.W.For the Alzheimer's Disease Neuroimaging, I., 2018. Amyloid involvement in subcortical regions predicts cognitive decline. Eur. J. Nucl. Med. Mol. Imaging 45, 2368–2376.

Choi, H., Lee, D.S., 2018. Generation of structural MR images from amyloid PET: application to MR-less quantification. J. Nucl. Med. 59, 1111–1117.

Choi, Y., Choi, M., Kim, M., Ha, J.-.W., Kim, S., Choo, J., 2018. Stargan: unified generative adversarial networks for multi-domain image-to-image translation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8789–8797.

Clark, C.M., Schneider, J.A., Bedell, B.J., Beach, T.G., Bilker, W.B., Mintun, M.A., Pontecorvo, M.J., Hefti, F., Carpenter, A.P., Flitter, M.L., 2011. Use of florbetapir-PET for imaging β-amyloid pathology. JAMA 305, 275–283.

Cohen, J.P., Luck, M., Honari, S., 2018. Distribution Matching Losses Can Hallucinate Features in Medical Image Translation. MICCAI. Springer, pp. 529–536.

Curtis, C., Gamez, J.E., Singh, U., Sadowsky, C.H., Villena, T., Sabbagh, M.N., Beach, T.G., Duara, R., Fleisher, A.S., Frey, K.A., 2015. Phase 3 trial of flutemetamol labeled with radioactive fluorine 18 imaging and neuritic plaque density. JAMA Neurol. 72, 287–294.

Frings, L., Hellwig, S., Spehl, T.S., Bormann, T., Buchert, R., Vach, W., Minkova, L., Heimbach, B., Kloppel, S., Meyer, P.T., 2015. Asymmetries of amyloid-beta burden and neuronal dysfunction are positively correlated in Alzheimer's disease. Brain 138, 3089–3099.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets. Adv. Neural Inf. Process. Syst. 2672–2680.

Hanseeuw, B.J., Betensky, R.A., Mormino, E.C., Schultz, A.P., Sepulcre, J., Becker, J.A., Jacobs, H.I.L., Buckley, R.F., LaPoint, M.R., Vannini, P., Donovan, N.J., Chhatwal, J.P., Marshall, G.A., Papp, K.V., Amariglio, R.E., Rentz, D.M., Sperling, R.A., Johnson, K.A., 2018. PET staging of amyloidosis using striatum. Alzheimers Dement. 14, 1281–1292.

Hegazy, M.A., Cho, M.H., Cho, M.H., Lee, S.Y., 2019. U-net based metal segmentation on projection domain for metal artifact reduction in dental CT. Biomed. Eng. Lett. 9, 375–385.

Herholz, K., Ebmeier, K., 2011. Clinical amyloid imaging in Alzheimer's disease. Lancet Neurol. 10, 667–670.

Huang, G., Liu, Z., Maaten, L.v.d., Weinberger, K.Q., 2016. Densely connected convolutional networks. arXiv:1608.06993.

Hwang, D., Kim, K.Y., Kang, S.K., Seo, S., Paeng, J.C., Lee, D.S., Lee, J.S., 2018. Improving the accuracy of simultaneously reconstructed activity and attenuation maps using deep learning. J. Nucl. Med. 59, 1624–1629.

Jagust, W.J., Landau, S.M., Koeppe, R.A., Reiman, E.M., Chen, K., Mathis, C.A., Price, J.C., Foster, N.L., Wang, A.Y., 2015. The Alzheimer's disease neuroimaging initiative 2 PET core: 2015. Alzheimer's Dement. 11, 757–771.

Kang, S.K., Seo, S., Shin, S.A., Byun, M.S., Lee, D.Y., Kim, Y.K., Lee, D.S., Lee, J.S., 2018. Adaptive template generation for amyloid PET using a deep learning approach. Hum. Brain Mapp. 39, 3769–3778.

Klunk, W.E., Engler, H., Nordberg, A., Wang, Y., Blomqvist, G., Holt, D.P., Bergström, M., Savitcheva, I., Huang, G.F., Estrada, S., 2004. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. Ann. Neurol. 55, 306–319.

Klunk, W.E., Koeppe, R.A., Price, J.C., Benzinger, T.L., Devous Sr, M.D., Jagust, W.J., Johnson, K.A., Mathis, C.A., Minhas, D., Pontecorvo, M.J., 2015. The Centiloid Project: standardizing quantitative amyloid plaque estimation by PET. Alzheimer's Dement. 11, 1–15 e14.

Landau, S.M., Breault, C., Joshi, A.D., Pontecorvo, M., Mathis, C.A., Jagust, W.J., Mintun, M.A., 2013. Amyloid-β imaging with Pittsburgh compound B and florbetapir: comparing radiotracers and quantification methods. J. Nucl. Med. 54, 70–77.

Lee, J.S., 2020. A review of deep learning-based approaches for attenuation correction in positron emission tomography. IEEE Trans. Radiat. Plasma Med. Sci. 1-1.

Leynes, A.P., Yang, J., Wiesinger, F., Kaushik, S.S., Shanbhag, D.D., Seo, Y., Hope, T.A., Larson, P.E., 2017. 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., jnumed 117, 198051.

McGraw, K.O., Wong, S.P., 1996. Forming inferences about some intraclass correlation coefficients. Psychol. Methods 1, 30.

Mesulam, M.M., Weintraub, S., Rogalski, E.J., Wieneke, C., Geula, C., Bigio, E.H., 2014. Asymmetry and heterogeneity of Alzheimer's and frontotemporal pathology in primary progressive aphasia. Brain 137, 1176–1192.

Miyato, T., Kataoka, T., Koyama, M., Yoshida, Y., 2018. Spectral normalization for generative adversarial networks. arXiv preprint arXiv:1802.05957.

Navitsky, M., Joshi, A.D., Kennedy, I., Klunk, W.E., Rowe, C.C., Wong, D.F., Pontecorvo, M.J., Mintun, M.A., Devous Sr, M.D., 2018. Standardization of amyloid quantitation with florbetapir standardized uptake value ratios to the Centiloid scale. Alzheimer's & Dementia 14, 1565–1571.

Nie, D., Cao, X., Gao, Y., Wang, L., Shen, D., 2016. Estimating CT Image from MRI Data Using 3D Fully Convolutional Networks. MICCAI. Springer, pp. 170–178.

Park, J., Hwang, D., Kim, K.Y., Kang, S.K., Kim, Y.K., Lee, J.S., 2018. Computed tomography super-resolution using deep convolutional neural network. Phys. Med. Biol. 63, 145011.

Prasoon, A., Petersen, K., Igel, C., Lauze, F., Dam, E., Nielsen, M., 2013. Deep Feature Learning For Knee Cartilage Segmentation Using a Triplanar Convolutional Neural Network. MICCAI. Springer, pp. 246–253.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional Networks for Biomedical Image Segmentation. MICCAI. Springer, pp. 234–241.

Sabri, O., Sabbagh, M.N., Seibyl, J., Barthel, H., Akatsu, H., Ouchi, Y., Senda, K., Murayama, S., Ishii, K., Takao, M., 2015. Florbetaben PET imaging to detect amyloid beta plaques in Alzheimer's disease: phase 3 study. Alzheimer's Dement. 11, 964–974.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273–289.

Villemagne, V.L., Mulligan, R.S., Pejoska, S., Ong, K., Jones, G., O'Keefe, G., Chan, J.G., Young, K., Tochon-Danguy, H., Masters, C.L., 2012. Comparison of 11 C-PiB and 18 F-florbetaben for Aβ imaging in ageing and Alzheimer's disease. Eur. J. Nucl. Med. Mol. Imaging 39, 983–989.

Wolk, D.A., Zhang, Z., Boudhar, S., Clark, C.M., Pontecorvo, M.J., Arnold, S.E., 2012. Amyloid imaging in Alzheimer's disease: comparison of florbetapir and Pittsburgh compound-B positron emission tomography. J. Neurol. Neurosurg. Psychiatry 83, 923–926.

Yi, X., Walia, E., Babyn, P., 2019. Generative adversarial network in medical imaging: a review. Med. Image Anal. 58, 101552.

Yonekura, A., Kawanaka, H., Prasath, V.B.S., Aronow, B.J., Takase, H., 2018. Automatic disease stage classification of glioblastoma multiforme histopathological images using deep convolutional neural network. Biomed. Eng. Lett. 8, 321–327.

Zhu, J.-.Y., Park, T., Isola, P., Efros, A.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2223–2232.

Wev reference

Kolibash, S. A., Centiloid Level-2 Analysis of [^{18}F]Florbetaben (FBB) and [^{18}F]Florbetapir (FBP) PET Image Data using the ADNI Pipeline. http://adni.loni.usc.edu/wp-content/themes/freshnews-dev-v2/documents/pet/ADNI%20Centiloids%20Final.pdf (accessed August 12 2020).