Convolutional Neural Network Application for Brain PET Image Translation to Support Alzheimer's Disease Diagnosis

Natacha Lambert, Daniel Roquet, Matéo Mahaut, Nicolas Bitouzé, Gaël Chételat, Abderrahim Elmoataz

Convolutional Neural Network Application for Brain PET Image Translation to Support Alzheimer's Disease Diagnosis

Natacha Lambert*, Daniel Roquet†, Matéo Mahaut‡, Nicolas Bitouzé§, Gaël Chételat¶ and Abderrahim Elmoataz‖
*Neuropresage Team & GREYC Laboratory, University of Caen Normandy, 14000 Caen, France
Email: natacha.lambert@unicaen.fr
† Neuropresage Team, INSERM, U1237, University of Caen Normandy, Cyceron, 14000 Caen, France
Email: roquet@cyceron.fr
‡Department of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Catalunya
Email: mateo.mahaut@upf.edu
§Forlabs, 14280 Saint-Contest, France
Email: nicolas.bitouze@forlabs.fr
¶ Neuropresage Team, INSERM, U1237, University of Caen Normandy, Cyceron, 14000 Caen, France
Email: chetelat@cyceron.fr
‖ Image Team, GREYC Laboratory, CNRS 6072, University of Caen Normandy, CEDEX 14050 Caen, France
Email: abderrahim.elmoataz-billah@unicaen.fr

***Abstract*—Alzheimer's Disease (AD) is supported by several complementary neuroimaging scans including ^{18}F-fluorodeoxyglucose PET (FDG-PET), ^{18}F-AV45 florbetapir PET (AV45-PET) and tau PET. Limiting PET scans would allow reducing cost and radiation exposure. Interestingly, the early PET acquisition of the AV45 scan would allow providing information similar to a FDG-PET scan. This study aims to perform an unidirectional translation from eAV45-PET to FDG-PET to optimize the similarities between both scan so that both complementary informations could be obtained from a single visit and radiotracer injection. The eAV45-PET to FDG-PET image translation is made by a ResNet neural network as well as a specific encoding and decoding algorithms. Results show good performances (0.00160 for Mean Square Error, 0.10301 for Normalized Root Mean Square Error, 33.80288 for Peak Signal to Noise Ratio and 33.80288 for Structural Similarity), however the technique still needs to be optimized to reduce the intensity discrepancy in some brain regions especially the frontal cortex.**

***Index Terms*—Alzheimer's disease, neuroimaging, PET scans, image generation, deep learning, translation**

I. Introduction

Alzheimer's Disease (AD) is the most common type of dementia and affects more than 35 million people in the world. This number tends to increase due to the aging of the population. There is no known cure for this disease, and the cause is not yet understood, but treatments can slow down its progression. Sadly the early and differential diagnosis of Alzheimer's disease (AD) is still challenging (M. McKhann and al. [1]).

Various parameters can be used to diagnose AD, in particular with neuroimaging. Several neuroimaging exams can be realized to see different markers of the disease: brain Magnetic Resonance Imaging (MRI) scan for hippocampal atrophy, and three different Positron Emission Tomography (PET) scans: ^{18}FF-fluorodeoxyglucose PET (FDG-PET), amyloid and tau PET to assess for metabolism changes, amyloid and tau PET for tau deposits, respectively. Although complementary, acquiring 3 PET acquisitions requires multiple injections, each being invasive and expensive (Vanhoutte and al. [2], R. Jack and al. [3]).

Amyloid AV45 is a reliable measure of amyloid deposition. Typically, the brain image is acquired at 50 to 70 minutes after injection, which is the necessary duration for maximal binding of the tracer onto amyloid deposits. Early acquisition of the AV45 tracer, between 0 to 10 minutes, is however related to the diffusion of the tracer within the brain vascular system before its binding onto the amyloid deposit, providing an image of the cerebral blood flow. This type of PET, named eAV45-PET (in opposition to the AV45-PET with the usual delay) shares a lot of similarities with FDG-PET, as expected when measuring cerebral blood flow and metabolism (Vanhoutte and al. [2], Asghar and al. [4], Hsiao and al. [5]). However, eAV45-PET seems less sensitive to AD than FDG-PET. From this observation, the aim of this article is to generate the result of a FDG-PET from an eAV45-PET, using a Convolutional Neural Network (CNN). Providing such a proxy of the FDG-PET would enable the procurement of both metabolism-proxy and amyloid map with only one injection, reducing acquisition cost and being less invasive. In addition, we aim to provide a synthesized metabolism map that is more sensitive to AD than the original eAV45-PET scan.

Our method, which uses a CNN combined with an encoder and a decoder algorithms, allows the generation of a FDG-PET image, named artificial FDG-PET, from an eAV45-PET

scan. At the best of our knowledge we are the first to propose such translation. This artificial FDG-PET image is close to the true FDG-PET scan, which correspond to the original, with stronger proximity than the initial eAV45-PET scan.

This paper is organized as follows: Section II proposes a review of some related work. Then, Section III details the methodology of this study, in terms of data acquisition and processing, neural network choice, and evaluation parameters. Afterward, Section IV presents the results of this study and Section V discuss them. Finally, Section VI concludes this paper with planned future works.

II. Related Work

eAV45-PET is a recent method with first reports dating from a decade ago: Hsiao and al. [5] point out the proximity between FDG-PET and eAV45-PET by showing a significant correlation between these two images. Asghar and al. [4] subsequently showed its relevance in the detection of cognitive disease, although with a lower sensitivity.

So far, eAV45-PET has been used as is, as "pseudo-FDG", but no proper synthesis of FDG-PET has been realised. Hazrat and al. [6] present a review on medical imaging generation and PET images are not the common choice, especially for eAV45-PET. An explanation for this is that deep learning, which is one of the best choice for image generation, requires a lot of data to provide reliable results. Unfortunately, eAV45-PET being quite recent, data are available only in a limited quantities. Due to these difficulties, eAV45-PET suffers from a lack of familiarity and use outside of the research circle, unlike the more widely used other PET or MRI scans.

Concerning image generation, although we do not use this method in this work, the most popular technology is Generative Adversarial Network (GAN). This technology is based on a training through a discriminator. This discriminator is also a neural network which updated dynamically to evaluate the accuracy of the first one. This model can be unidirectional or bidirectional. Some very recent image generation works use the stable diffusion method (Rombach and al. [7]). Nonetheless, to our knowledge, they have not been scaled to PET scans, and cannot yet be applied to small data set like eAV45-PET. The efficiency of GAN makes them popular in many fields including the medical field (Hazrat and al. [6]). In our case, translation is the type of transformation of interest, as it can generate a new image (artificial FDG-PET) from an existing image (eAV45-PET). In the literature, the general case is to generate an MRI image from another MRI, Computed Tomography (CT) or PET modality. Due to the complexity and time-consuming calculations, image generation is generally approached in 2D, even for medical 3D images with the use of slices. Lots of popular GAN networks like CycleGAN or pix2pix (Zhu and al. [8]) show good results with 2D medical images. For example, Tang and al. [9] adapt a pix2pix GAN model, with Unet generator, to synthesis 2D CT with MRI scan. Likewise, Unsupervised Attention Guided GAN (uagGAN) is used in Abu-Srhan and al. [10] to make a 2D bidirectional translation between MRI and CT. This network needs to have the same shape between the two images, which is not a problem for brain images, and shows better results than other GAN networks. Gu and Zheng [11] use GANs combinaison Dual3D&PatchGAN based on fuzz-c-mean algorithm, to generate CT from MRI by reconstructing 3D scans with 2D slices. But CT is not the only modality synthesized: Marco Conte and al. [12] use 2D GAN model to synthesis missing T1 MRI or FLAIR (FLuid Attenuated Inversion Recovery) MRI from post-contrast T1 MRI or T2 MRI. The same idea is taken up by Dai and al. [13], which propose a derived of StarGAN to create an unified GAN which can generate 3D multimodal MRI like T1, T2 or even FLAIR depending on the missing data.

For PET generation, most of the work is in 3D. Zhang and al. [14] introduce BPGAN, based on BicyleGAN, an unidirectional network which can transform MRI to PET and these PET scans show good results for Alzheimer Disease diagnosis. Other studies propose bidirectional GAN, to pass MRI to PET and vice versa. Hu and al. [15], for example, use DenseNet and U-Net networks to construct their model named BMGAN which can pass one modality to the other. In an other hand, Lin and al. [16] build 3D RevGAN, based on RevGAN and SRCNN, which also performs well for the translation between MRI and PET images. We compare our work to this GAN state of the art in the results section.

Although GAN is the most popular method, it is not the only solution for image generation. Convolutional Neural Network (CNN) can also be used to perform image reconstruction. For example, Han [17] uses MRI image to obtain Computed Tomography (CT) with a DCNN model, or to reconstruct other MRI modalities like FLAIR with a U-Net model like Osman and M. Tamam [18], with good results. CNNs are easier to train than GANs with fewer data points, we have therefore chosen to focus on them.

III. Methodology

A. Data Population

To train our model, we need both FDG-PET and eAV45-PET scans, which are not the most common data. We use three different databases to maximize the number of available images.

First, 118 pairs of images (FDG-PET and eAV45-PET) were selected from the 'Imagerie Multimodale de la Maladie d'Alzheimer à un stade Precoce' (IMAP) Study (Caen, France). This protocol contains healthy participants with various ages, but most of them older than 60, and participants with different stages of cognition decline. Out of the 118 selected pairs of images, 45 belong to participants with no cognition problem (CN: Cognitive Neutral), 23 to participants with complaints about their memory but no clinical sign (SCD: Subjective Cognitive Decline), 32 to participants with Mild Cognitive Impairment (MCI) and 18 to participants with Alzheimer Dementia (AD). Inclusion and exclusion criteria are detailed in Mutlu and al. [19] and Wirth and al. [20]. For each subject, there may be several available image captures, taken

during different visits. We keep up to three of them, ignoring low-quality ones.

Second, 223 pairs of images were selected from the Age-Well randomized controlled trial of the Medit-Ageing European Project (Poisnel and al. [21]), which contains only healthy old adults. For the same reasons as before, for each subject, there may be several available image captures taken during different visits.

Third and last, 84 pairs of images were selected from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database[1], which are available in open access. Out of the 84 selected pairs of images, 16 belong to CN participants, 34 to SCD participants, 23 to MCI participants and 11 to AD participants. Only one visit per participants was included.

Participant demographics for each database are summarized in TABLE I. The breakdown is not given participant by participant but pair of images by pair of images.

The compound database is divided in two subgroups. 260 pairs of PET images are used for the training dataset (61% of the database) and 165 pairs of PET images for testing the model (39% of the database). The separation is made in such a way that each source database, gender and pathology, are equally represented in both training and testing datasets. It also guarantees that if a subject has more than one visit, all its visits are in the same dataset. Once all the criteria are fulfilled, the distribution is made randomly. In the first dataset, we use 80% (208 pairs of images) for training and 20% (52 pairs of images) for validation.

TABLE I
PARTICIPANTS DEMOGRAPHICS (BY PAIR OF IMAGES)
(AD: ALZHEIMER DISEASE, MCI: MILD COGNITIVE IMPAIRMENT, SCD: SUBJECTIVE COGNITIVE DECLINE, CN: COGNITIVE NEUTRAL)

Characteristics	**Agewell** n=223	**IMAP** n=118	**ADNI** n=84	**Total** n=425
F gender, n (%)	139 (62,3)	55 (46,6)	43 (51,2)	237 (55,8)
AD, n (%)	0	18 (15,3)	11 (13,1)	29 (6,8)
MCI, n (%)	0	32 (27,1)	23 (27,4)	55 (12,9)
SCD, n (%)	0	23 (19,5)	34 (40,5)	57 (13,4)
CN, n (%)	223 (100)	45 (38,1)	16 (19,0)	284 (66,8)
Age, years (std)	$70 \pm 3,8$	$65,7 \pm 15,1$	$73 \pm 7,2$	$69,4 \pm 9,3$
Education, years (std)	$13 \pm 3,1$	$12,5 \pm 3,1$	$16,46 \pm 2,7$	$13,5 \pm 3,4$

B. Data Acquisition and Processing

IMAP and Age-well databases share the same acquisition parameters, but differ from ADNI ones (the parameters are available on its website). Image processing was however common across the three databases, and has already been published by Vanhoutte et al. [2]. All the data processing is made with Statistical Parametric Mapping 12 (SPM12) software[2] and a quality evaluation is performed to only use good quality images.

1) MRI: MRI images are not used directly for the neural network but are necessary for PET image processing. First for the co-registration onto the MRI T1 image, then for the gray matter segmentation of the PET images like in the usually applied diagnosis protocols. T1-weighted imaging parameters were all acquired from 3T scanners, although from various manufacturers (Philips, Siemens). After acquisition, each MRI image is normalized to the Montreal Neurological Institute (MNI) space[3] and segmented into gray matter, white matter and cerebrospinal fluid. Resulting images have the following parameters: matrix size = $121 \times 145 \times 121$, voxel size = $1.5mm \times 1.5mm \times 1.5$mm.

2) PET: For FDG-PET scans, 5 mCi of ^{18}F-fluorodeoxyglucose are injected, then 3D images are dynamically taken as follows:

- Between fifty minutes and sixty minutes after injection with a ten-minutes long acquisition for Age-well and IMAP,
- Between thirty and sixty minutes after injection with six five-minutes long acquisition for ADNI.

For eAV45-PET scans, 10 mCi of ^{18}F-AV45 florbetapir are injected and 3D images are dynamically taken right after:

- Every 60s for 10 minutes, resulting in 10 images, for IMAP and Age-well,
- Over a total duration of 16 minutes divided in 16 different time lapses (every fifteen seconds for the first minute, every thirty seconds until the third minute, every minute until the sixth minute, every two minutes until twelfth minute and the last on the sixteenth minute) for ADNI.

Times-frames from 2 to 5 min for eAV45-PET scan were used in the subsequent analysis. Then, we combine the images obtained during each injection into a single image by taking their arithmetic mean. Each mean PET image is next co-registered onto their corresponding MRI. After that, all 3D PET images are normalized to the MNI space, and therefore shared the same matrice size and voxel resolution as the MRI images (matrix size = $121 \times 145 \times 121$, voxel size = $1.5mm \times 1.5mm \times 1.5$mm). Eventually, to allow inter-subject comparison and to have comparable voxel values, all PET data are scaled (intensity normalization) by the mean value of the cerebellum for the FDG-PET and the brainstem for the eAV45-PET.

C. Neural Network Methods

Image generation is usually performed by Generative Adversarial Network (GAN), whether for generation, reconstruction or, in our case, translation. Nevertheless, these neural networks are mostly used for 2D images. Despite some studies used derived 2D for 3D images (Hu and al. [15], Lin and al. [16], Zhang and al. [14]), this approach is particularly time-consuming on development and training, and requires an important dataset. With this context, and because our PET images are scaled, we choose to use a 3D Convolutional Neural Network to perform our image translation, like previous

[1]https://adni.loni.usc.edu/

[2]https://www.fil.ion.ucl.ac.uk/spm/software/spm12/

[3]https://www.bic.mni.mcgill.ca/ServicesAtlases/ICBM152NLin2009

studies on MRI, CT and FLAIR (Han [17], Osman and Tamam [18]).

We also use Monai [22], an open-source Python library that contains the most recognized neural network adapted for 2D and 3D medical imaging. It also guarantees the correct treatment of PET and MRI images, which have a special format alias: nifti. After some preliminary tests on different networks, like Unet or Resnet, we select a derived ResNet network called SegResNet [23], with Dice Coefficient for validation measure, batch size=4, epochs=1000, and Adam optimizer. Fig. 1 summarizes the network architecture and describes the different convolutions applied by the neural network.

PET images having their intensity being normalized, with intensity ranging between 0 and approximately 2, voxels values can therefore be classified into a limited set of classes, in agreement with SegResNet requirement. Preliminary tests were conducted to define the optimal number of classes and their range of values. An equal difference between all classes is the easiest choice but shows poor results, especially because high intensity is rare. Fig. 2 shows a classic distribution of voxel values for a scaled FDG-PET scan. Removing the voxels from 0 to 0.3, which correspond to the background in neuroimaging, the majority of values are aggregated between 0.3 and 1.5. That is why we need more detail between these two values. Likewise, choosing as many classes as possible shows disappointing results. Indeed, a PET image has small values and choosing too many classes results in recognizing very few of them because of their proximity. The best choice is to use eight classes, with 0 for the background and 7 for higher intensity, with this separation:

- values below 0.3 belong to the background class 0 (because such values are discarded in the post-processing step),
- values between 0.3 and 1.5 are spread into 6 classes (numbered from 1 to 6) of equal width 0.2
- values above 1.5 belong to the last class (7).

An encoding algorithm is made to transform each PET image to an image with the same dimensions but with discrete values in $0, \ldots, 7$, before passing it to the SegResNet Network.

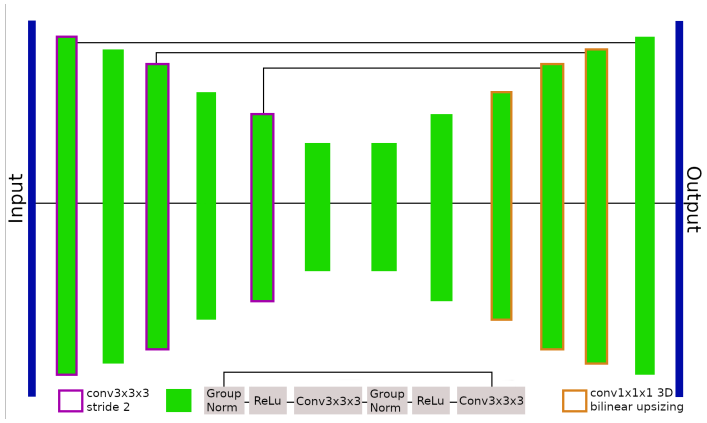


Fig. 1. SegResNet architecture described in Myronenko [23]: all steps are left to right, with ResNet block

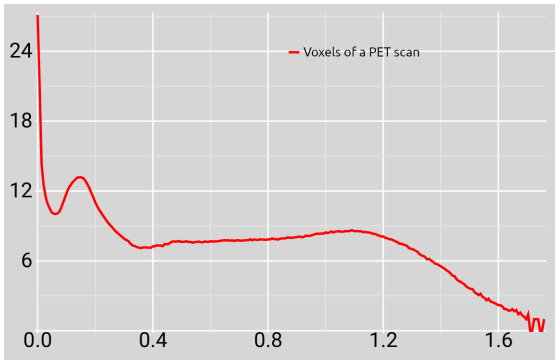


Fig. 2. Intensity histogram of FDG-PET image: X axis is for values and Y axis for the percentage of voxels having this value

Once the model is trained, PET images are reconstructed to have a range of values that is more appropriate for medical analyses. For that, a decoding algorithm based on network probabilities has been created. For each FDG-PET image created by the SegResNet model, we also save a probability map for each class, represented by a 4D images with (121, 145, 121, 8) for size. We also use the eAV45-PET scan to estimate where the reconstructed value should lie within the range assigned to its class. The probability map value is used to weigh this observation: we consider that a high probability means that a voxel has more chance to be a middle value of its class and vice versa. Therefore, to assign a value to a given voxel in the artificial FDG-PET image, we initially assign either the highest or the lowest value of its class by considering whether its eAV45-PET counterpart voxel has a high or low value. We then correct the result by shifting it towards the middle of the class, by an amount which is proportional to the probability of the class in the probability map, as shown in Equation 1. One can note that classes of interest have a width of 0.2, therefore if such a class has probability 1, the formula returns the mid value of the class.

$$\text{Voxel value} = \begin{cases} \text{Low Class value} + \frac{p}{10} & \text{if } EarlyAV45 \text{ low,} \\ \text{High Class value} - \frac{p}{10} & \text{if } EarlyAV45 \text{ high,} \end{cases} \quad (1)$$

where p is the probability that the voxel belongs to its class.

In the end, we have our artificial FDG-PET image. However, for medical interpretation, only gray matter is relevant for this modality. Accordingly, artificial FDG-PET images are also segmented for gray matter using the MRI T1 gray matter probability map on SPM12, with Equation 2.

$$\text{output}[v] = \begin{cases} \text{FDG-PET}[v] & \text{if GM}[v] > \max \begin{pmatrix} \text{WM}[v] \\ \text{CL}[v] \\ 0.3 \end{pmatrix}, \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

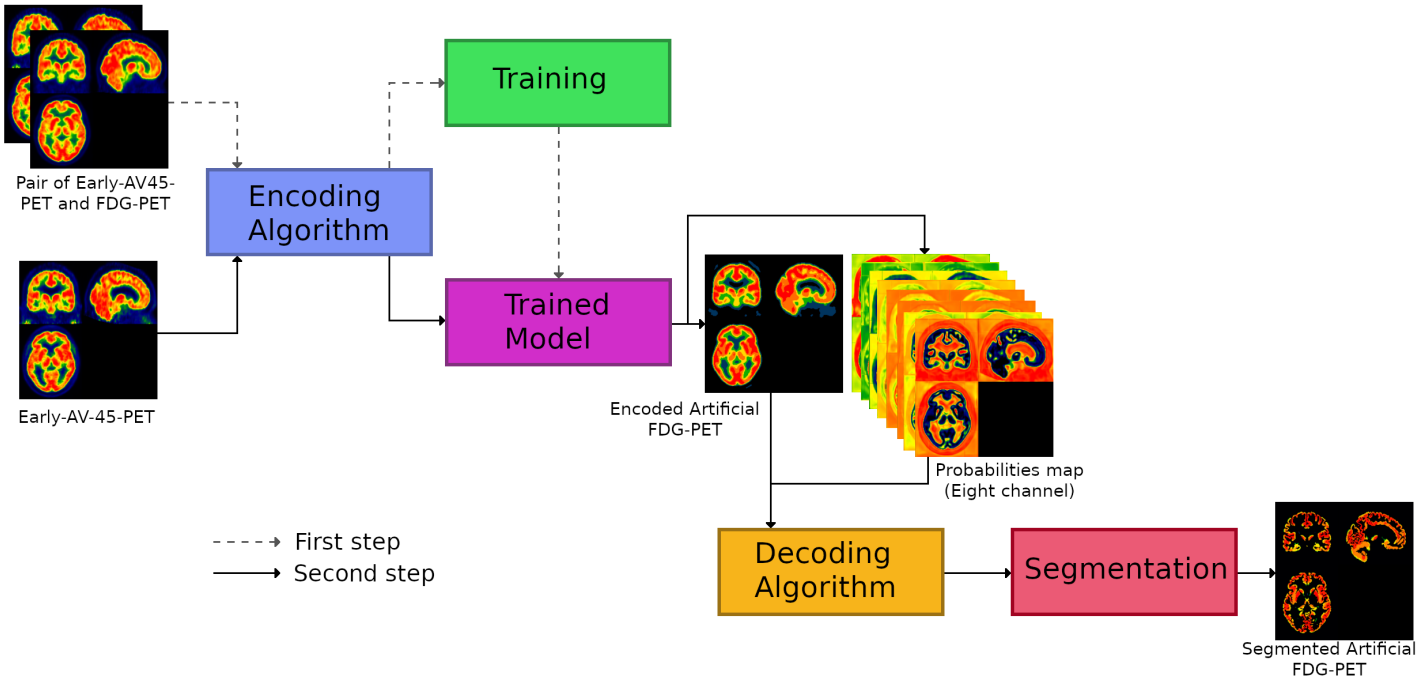


Fig. 3. Complete FDG-PET generation process

where output is the segmented FDG-PET, v is voxel value, GM is Gray Matter, WM is White Matter and CL is cerebrospinal liquid.

Fig. 3 summarizes the whole process. First, we have a training phase, with SegResNet model, our 260 pairs of PET images (80% for training, 20% for validation) encoded with an eight class encoder. Then, the testing phase begins, with our 165 other eAV45-PET scans encoded to the trained model that gives us 165 artificial FDG-PET images, which are decoded with the decoding algorithm and segmented to only keep gray matter. The final step is the evaluation of artificial FDG-PET images, in comparison with corresponding true FDG-PET scans.

D. Evaluation Metrics

To evaluate the model, we use several metrics to estimate the similarity between artificial FDG-PET images and true FDG-PET scans. For that, common measures in the literature were chosen: Mean Square Error (MSE) and Normalized Root Mean Square Error (NRMSE) [24] to compute errors between the artificial and true FDG-PET, combined with Peak Signal to Noise Ratio (PSNR) [25] for the quality of the image and Structural Similarity (SSIM) [26] to estimate proximity between the artificial and true FDG-PET images.

First, we aim to assess the quality of the artificial FDG-PET generation by comparing it to the true FDG-PET scan. Then, we estimate whether artificial FDG-PET has a better similarity to true FDG-PET than eAV45-PET, considered in previous studies as pseudo-FDG-PET (Asghar and al. [4] and Vanhoutte and al. [2]). Last, we compare our results with studies (Zhang and al. [14], Hu and al. [15], Lin and al. [16]) that also performed an artificial FDG-PET image, although the image used for synthesis is not an eAV45-PET but MRI scan.

E. Experimental Set-Up

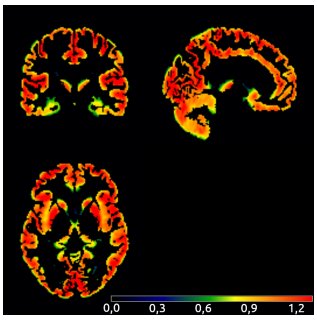
All algorithms (except data processing, which is done with SPM12 as previously mentioned) use Python. Pytorch and Monai [22] are required for neural network, sklearn and skimage for evaluation measure, with utility libraries: numpy, pandas and nibabel. Computations are made on an internal computing platform with 2 Intel Xeon E5-2640 v4 2.40GHz processors, 40 cores, 256Go RAM and 8 MSI GeForce GTX 1080 Ti 3584 Cores 11G RAM for GPU. A single GPU is required to execute the model and the image reconstruction.

IV. Results

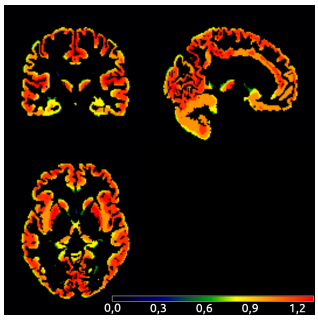
A. Comparison with true FDG-PET

The aim of this study is to generate an artificial FDG-PET image from an eAV45-PET scan that maximizes the similarity with the true-FDG-PET scan as shown in Fig. 4. Artificial FDG-PET image (4(b)) is very similar to the corresponding FDG-PET image (4(a)). We observe some smoothing differences between two different intensity areas, due to the model and decoding algorithm, but intensity areas are the same on the same brain region. Artificial FDG-PET is slightly less high-valued than the corresponding true FDG-PET, especially in frontal and occipital brain areas, but we obtain a very similar result.

To confirm the good quality of the image, TABLE II shows their computed evaluation measures. Each value is calculated between artificial and true FDG-PET images. We separate PET images into two categories: with and without pathology,



(a) True FDG-PET



(b) Artificial FDG-PET

Fig. 4. Comparison between corresponding True 4(a) and Artificial 4(b) FDG-PET on a healthy subject (CN)

TABLE II
RESULTS OF EVALUATION MEASURES BETWEEN ARTIFICIAL AND TRUE FDG-PET IMAGES (MEAN SQUARE ERROR, NORMALIZED ROOT MEAN SQUARE ERROR, PEAK SIGNAL TO NOISE RATIO AND STRUCTURAL SIMILARITY)

	General Results	**CN Results**	**Patient Results**
MSE	0.00160 ± 0.00021	0.00163 ± 0.00029	0.00152 ± 0.00026
NRMSE	0.10301 ± 0.00521	0.10222 ± 0.00651	0.10467 ± 0.00884
PSNR	33.80288 ± 0.42296	33.95528 ± 0.54451	33.48081 ± 0.62873
SSIM	0.99528 ± 0.00045	0.99539 ± 0.00056	0.99505 ± 0.00077

as FDG-PET aims to be used for clinical diagnosis in the future. Artificial images obtain good results. For SSIM, we reach 0.99528, which a very high score for this measure whose maximum value is 1. Likewise, we achieve 33.80288 for PSNR. As for the two error values, we have 0.00160 for MSE and 0.10301 for NRMSE, which lower is better for both measures, with 0 as unreachable bound.

We also notice that CN and patients scores are similar. SSIM does not differ, with 0.99539 and 0.99505 respectively. PSNR varies by 1% between CN and patients results. As for the error, MSE varies by 8% and NRMSE by 2%. Except for MSE, the others measures is in favor for CN results but the variance is very small, which is a good point.

B. Comparison with eAV45-PET

In order to measure the quality of our work, we compared the similarity of artificial FDG-PET and eAV45-PET to true FDG-PET images in TABLE III. Although eAV45-PET has good results, our processing allowed to reduce MSE by 36% and reduce NRMSE by two points. PSNR climbs two points, which illustrates improvement of artificial FDG-PET over eAV45-PET. SSIM does not show a great difference between eAV45-PET and artificial FDG-PET, with 0.99337 and 0.99528 respectively. Nevertheless, this can be explained by the structural similarity of the PET images to each other.

C. Comparision with previous studies

To our knowledge, no study has generated artificial FDG-PET images from eAV45-PET scans. Therefore, we were only able to compare our study with MRI-based artificial FDG-PET studies. Our results reached their best scores on the evaluation metrics when computed on segmented FDG-PET (see in tablename II). However, since some of the previous studies computed the image on the whole unsegemented brain (Zhang and al. [14], Hu and al. [15], Lin and al. [16]), we also assessed our metrics on unsegmented images in TABLE IV. Results showed lower similarity when computed in the entire brain (33.80288 to 31.08430 for PSNR and 0.99528 to 0.95479 for SSIM) but were still better than previous studies. When they reached 25.08, 27.88117 and 29.42 for PSNR, we achieved 31.08430. This correspond to an increase from 6% to 24%. In the same way, previous studies performed 0.6646, 0.89008 and 0.8176 for SSIM while we had 0.95479. We also improved this score by 7% to 44%.

V. LIMITATIONS

While results are promising, our work have some limitations.

First of all, as we previously said, there are no other works generating artificial FDG-PET image from eAV45-PET scan. Therefore, we compared our results with studies (Zhang and al. [14], Hu and al. [15], Lin and al. [16]) with MRI-based artificial FDFG-PET. Moreover, GAN network used in comparative studies are bidirectional which means that they can also recreate an MRI image from an FDG-PET scan. Our network is unidirectional and can only recreate an FDG-PET from an eAV45-PET. Nevertheless, the reverse operation has no interest in this case, because eAV45-PET is not used for diagnosis.

Then, our method generates an artificial FDG-PET image with a convolutional neural network, less greedy both in terms of computation time and data than a GAN. Although the result is better than previous studies (Zhang and al. [14], Hu and al. [15], Lin and al. [16]), our method does not generate a perfect FDG-PET image, such as the smoothing problems previously mentioned (Fig. 4). Our compounded database contains only 425 pairs for images, which is already sufficient to have a suitable result. However, reinforcing it by adding images, in particular of MCI and AD patients, could improve the results. Upgrading our encoder and decoder algorithms can also smooth intermediates values.

Another limitation is about good results of the eAV45-PET scan (TABLE III). Although we improve the results on all scores, artificial FDG-PET quality can be partially explained by eAV45-PET quality. Although the improvement of artificial FDG-PET over eAV45-PET is clearly present, it remains

TABLE III
COMPARISON BETWEEN EAV45-PET AND ARTIFICIAL FDG-PET WITH TRUE FDG-PET

	eAV45-PET	**Artificial FDG-PET**
MSE	0.00251 ± 0.00030	0.00160 ± 0.00021
NRMSE	0.12941 ± 0.00692	0.10301 ± 0.00521
PSNR	31.83984 ± 0.48960	33.80288 ± 0.42296
SSIM	0.99337 ± 0.00062	0.99528 ± 0.00045

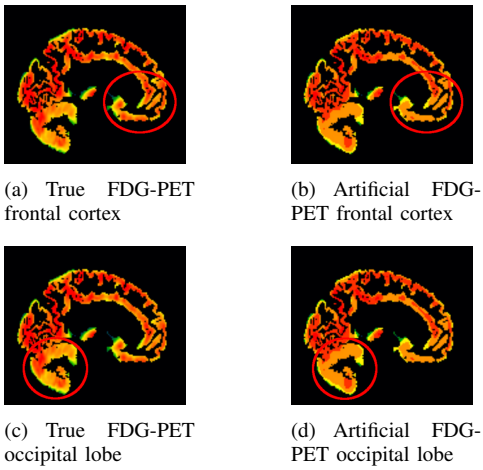
TABLE IV
COMPARISON WITH SIMILAR STUDIES ON EVALUATION MEASURES BETWEEN TRUE AND ARTIFICIAL FDG-PET ON THE WHOLE BRAIN (PEAK SIGNAL TO NOISE RATIO AND STRUCTURAL SIMILARITY)

	BPGAN [14]	**BMGAN [15]**	**3D RevGAN [16]**	**SegResNet**
PSNR	25.08	27.88117	29.42	31.08430
SSIM	0.6646	0.89008	0.8176	0.95479

moderate and it seems possible to improve the scores by a few points.

At last, we see on Fig. 4 that some brain regions, like frontal cortex or occipital lobe, have some discrepancies. Fig. 5 highlights the regions suffering from this concern. We observe that values on the true FDG-PET image 5(a) 5(c) are higher than artificial FDG-PET 5(b) 5(d), particularly in the frontal cortex 5(a) 5(b) and occipital lobes 5(c) 5(d) as shown by warmer colors in the true FDG-PET. This discrepancy can be found in eAV45-PET too, where it is even more marked. Although artificial PET-FDG images show good results (II) and stay close to true FDG-PET scans, we need to overcome these discrepancies by improving our algorithms. The challenge is that these discrepancies is not linear and depends on some parameters like pathology, which cannot be used as an input parameter as it would introduce a bias.

Future works will concentrate on improving our result, especially to erase discrepancies found on frontal cortex and occipital lobe. First of all, increasing the database seems complicated due to the limited quantity of available eAV45-PET scans. However, a solution to improve our method would be to enhance the encoder and decoder algorithms to better support the convolutional neural network. Then, we will post-process our artificial FDG-PET images to avoid the discrepancies. Mathematical models will be studied to overcome the discrepancies, without altering the brain regions which have a good convergence between the artificial FDG-PET and the true FDG-PET images. We will also focus further on intra group comparisons to better understand the discrepancies, and overcome them.



(a) True FDG-PET frontal cortex

(b) Artificial FDG-PET frontal cortex

(c) True FDG-PET occipital lobe

(d) Artificial FDG-PET occipital lobe

Fig. 5. Comparison between frontal cortex from True 5(a) and Artificial 5(b) FDG-PET and occipital lobe from True 5(c) and Artificial 5(d) FDG-PET on a healthy subject (CN)

VI. CONCLUSIONS

This study presents an innovative deep learning approach to generate an artificial FDG-PET images based on eAV45-PET scans, using convolutional neural network and an encoding and a decoding algorithms. This provides an artificial metabolism map in addition to the amyloid deposition map, which are both relevant for the diagnosis of Alzheimer's disease, at no extra cost and without any extra radioactive injection. Computed artificial FDG-PET images show high similarity with true FDG-PET scans (0.00160 for MSE, 0.10301 for NRMSE, 33.80288 for PSNR and 33.80288 for SSIM), with better performances than previous MRI-based artificial FDG-PET or eAV45-PET images. However, some discrepancies in specific brain regions (frontal cortex and occipital lobe) have to be overcome with further developments.

ACKNOWLEDGMENT

We would like to warmly thank the patients and healthy volunteers of the IMAP and Age-Well studies, as well as the ADNI study which is not internal to our lab. We also thank A. Abbas, C. André, T. Anquetil, E. Arenaza-Urquijo, L. Barre, JC. Baron, A. Bejanin, P. Champetier, L. Chauveau, A. Chocat, A. Cognet, J. Dayan, R. De Flores, M. Delarue, N. Delcroix, B Desgranges, S. Egret, R. El Sadawy, F. Eustache, S. Fauvel, M. Fouquet, M. Gaubert, J. Gonneaud, D. Guilloteau, A. Hamel, S. Haudry, O. Hebert, T. Köbe, E. Kuhn, R. La Joie, B. Landeau, M. Leblond, V. Lefranc, A. Lehodey, A. Manrique, F. Mezenge, K. Mevel, I. Moulinet, J. Mutlu, V. Ourry, L. Paly, C. Palix, G. Poisnel, A. Pelerin, A. Perrotin, A. Quillard, G. Rauchs, C. Schupp, S. Sherif, C. Tomadesso, E. Touron, A.L. Turpin, F. Viader, N. Villain, and the Cyceron staff members for the administrative support and their help with the data acquisition. In more detail, we would like to thank F. Mezenge and B. Landeau for their help on data processing, L. Chauveau and R. Deflores for their knowledge of ADNI database and S. Haudry for his help on PET images. Last, we thank T. Lambert for proofreading.

REFERENCES

[1] G. M. McKhann and al., "The diagnosis of dementia due to alzheimer's disease: Recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease," *Alzheimer's & Dementia*, vol. 7, no. 3, pp. 263–269, 2011.

[2] M. Vanhoutte and al., "Evaluation of the early-phase [18f]av45 pet as an optimal surrogate of [18f]fdg pet in ageing and alzheimer's clinical syndrome," *NeuroImage: Clinical*, vol. 31, p. 102750, 2021.

[3] C. R. Jack and al., "Nia-aa research framework: Toward a biological definition of alzheimer's disease," *Alzheimer's & Dementia*, vol. 14, no. 4, pp. 535–562, 2018.

[4] M. Asghar and al., "Dual-phase [18f]florbetapir in frontotemporal dementia," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 46, pp. 304–311, 2019.

[5] I.-T. Hsiao and al., "Correlation of early-phase 18f-florbetapir (av-45/amyvid) pet images to fdg images: Preliminary studies: Reply to ballinger," *European journal of nuclear medicine and molecular imaging*, vol. 39, p. 1513, 2012.

[6] A. Hazrat and al., "The role of generative adversarial networks in brain mri: a scoping review," *Insights into Imaging*, vol. 13, p. 98, 2022.

[7] R. Rombach and al., "High-resolution image synthesis with latent diffusion models," *arXiv*, 2021.

[8] J.-Y. Zhu and al., "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Computer Vision (ICCV), 2017 IEEE International Conference on*, 2017, pp. 2242–2251.

[9] B. Tang and al., "Dosimetric evaluation of synthetic ct image generated using a neural network for mr-only brain radiotherapy," *Journal of Applied Clinical Medical Physics*, vol. 22, no. 3, pp. 55–62, 2021.

[10] A. Abu-Srhan and al., "Paired-unpaired unsupervised attention guided gan with transfer learning for bidirectional brain mr-ct synthesis," *Computers in Biology and Medicine*, vol. 136, p. 104763, 2021.

[11] Y. Gu and Q. Zheng, "A transfer deep generative adversarial network model to synthetic brain ct generation from mr images," *Wireless Communications and Mobile Computing*, vol. 2021, 2021.

[12] G. M. Conte and al., "Generative adversarial networks to synthesize missing t1 and flair mri sequences for use in a multisequence brain tumor segmentation model," *Radiology*, vol. 299, no. 2, pp. 313–323, 2021.

[13] X. Dai and al., "Multimodal mri synthesis using unified generative adversarial networks," *Medical Physics*, vol. 47, no. 12, pp. 6343–6354, 2020.

[14] J. Zhang and al., "Bpgan: Brain pet synthesis from mri using generative adversarial network for multi-modal alzheimer's disease diagnosis," *Computer Methods and Programs in Biomedicine*, vol. 217, p. 106676, 2022.

[15] S. Hu and al., "Bidirectional mapping generative adversarial networks for brain mr to pet synthesis," *IEEE Transactions on Medical Imaging*, vol. 41, no. 1, pp. 145–157, 2020.

[16] W. Lin and al., "Bidirectional mapping of brain mri and pet with 3d reversible gan for the diagnosis of alzheimer's disease," *Frontiers in Neuroscience*, vol. 15, p. 646013, 2021.

[17] X. Han, "Mr-based synthetic ct generation using a deep convolutional neural network method," *Medical Physics*, vol. 44, no. 4, pp. 1408–1419, 2017.

[18] A. F. I. Osman and N. M. Tamam, "Deep learning-based convolutional neural network for intramodality brain mri synthesis," *Journal of Applied Clinical Medical Physics*, vol. 23, no. 4, p. e13530, 2022.

[19] J. Mutlu and al., "Distinct influence of specific versus global connectivity on the different Alzheimer's disease biomarkers," *Brain*, vol. 140, no. 12, pp. 3317–3328, 2017.

[20] M. Wirth and al., "Regional patterns of gray matter volume, hypometabolism, and beta-amyloid in groups at risk of alzheimer's disease," *Neurobiology of Aging*, vol. 63, pp. 140–151, 2018.

[21] G. Poisnel and al., "The age-well randomized controlled trial of the medit-ageing european project: Effect of meditation or foreign language training on brain and mental health in older adults," *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, vol. 4, no. 1, pp. 714–723, 2018.

[22] M. J. Cardoso and al., "Monai: An open-source framework for deep learning in healthcare," *arXiv-CS-Machine Learning*, 2022.

[23] A. Myronenko, "3d mri brain tumor segmentation using autoencoder regularization," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Springer International Publishing, 2019, pp. 311–320.

[24] A. Botchkarev, "A new typology design of performance metrics to measure errors in machine learning regression algorithms," *Interdisciplinary Journal of Information, Knowledge, and Management*, vol. 14, pp. 45–76, 2019.

[25] J. F. de Boer and al., "Improved signal-to-noise ratio in spectral-domain compared with time-domain optical coherence tomography," *Opt. Lett.*, vol. 28, no. 21, pp. 2067–2069, 2003.

[26] Z. Wang and al., "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.