

Diffused Multi-scale Generative Adversarial Network for low- dose PET images reconstruction

Xiang Yu

Zhejiang University

Dao Yan Hu

The College of Biomedical Engineering and Instrument Science of Zhejiang University

Qiong Yao

The Second Affiliated Hospital of Zhejiang University School of Medicine

Yu Fu

Zhejiang University

Yan Zhong

The Second Affiliated Hospital of Zhejiang University School of Medicine

Jing Wang

The Second Affiliated Hospital of Zhejiang University School of Medicine

Mei Tian

Fudan University

Hong Zhang

hzhang21@zju.edu.cn

The College of Biomedical Engineering and Instrument Science of Zhejiang University

Research Article

Keywords: Deep learning, Positron Emission Tomography, Low-dose PET, Image reconstruction

Posted Date: December 9th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-5151497/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BioMedical Engineering OnLine on February 9th, 2025. See the published version at <https://doi.org/10.1186/s12938-025-01348-x>.

Abstract

Purpose: The aim of this study is to reconstruct low-dose PET (L-PET) images to full-dose PET (F-PET) images based on our Diffused Multi-scale Generative Adversarial Network (DMGAN) to resolve the trade-off between lowering radiation exposure and sustaining diagnostic accuracy.

Methods: The proposed method including two modules: the diffusion generator and the u-net discriminator. The goal of the first module is to get different information from different levels, enhancing the generalization capability of the generator to the image and improving the stability of the training. Generated images are inputted into the u-net discriminator, extracting details from both overall and specific perspectives to enhance the quality of the reconstructed F-PET images. We conducted evaluations encompassing both qualitative assessments and quantitative measures. In terms of quantitative comparisons, we employed two metrics, structure similarity index measure (SSIM) and peak signal-to-noise ratio (PSNR) to assess the performance of diverse methods.

Results: The proposed approach demonstrated the highest PSNR and SSIM values outperforming all compared methods, which improved PSNR by at least 6.2% over the other methods. Compared to other methods, the synthesized full-dose PET image generated by our method exhibits a more precise voxel-level metabolic intensity distribution, leading to a sharper and clearer visualization of the epileptic focus.

Conclusions: The proposed method can restore original details from low-dose PET images, surpassing the performance of other models trained on identical datasets, which can resolve the conflict between minimizing radiation exposure and sustaining diagnostic accuracy.

Introduction

As one of the most commonly utilized medical imaging techniques, PET plays a key role in navigated surgery, medical assessment and clinical examination[1], which detects biochemical and physiological changes unlike other imaging technologies, such as magnetic resonance imaging and computed tomography[2]. Because biochemical and physiological alterations frequently precede anatomical changes, PET is also extensively employed for preventive treatment and early disease identification. PET can assess molecular changes within the human body in vivo evaluations. In spite of the considerable advantages offered by PET, an increasing concern among individuals pertains to the potential health hazards linked to exposure to radiation during the scanning procedure. For example, the administered activity is frequently limited due to concerns regarding radiation exposure to patients in clinical practice, because more radiation dose has the potential to raise the cancer risk and harm the body to some extent[3][4]. Hence, the utilization of L-PET images, enabling image acquisition with minimal radiation exposure, has garnered significant interest among researchers[5][6]. Nevertheless, the L-PET images display increased noise, reduced contrast, and more pronounced artifacts compared to the F-PET images[5], this renders them challenging for physicians to perform a precise diagnosis. Therefore, it has strong practical significance to obtain high-quality images from low-dose images to minimize image exposure while preserving image quality.

Various techniques have been introduced to improve the quality of PET images. One approach for achieving high-quality PET images involves integrating prior information into the image reconstruction process[7]. This method allows for the direct incorporation of imaging physics information. However, it faces challenges related to intensive computation, and access to the physics projection model is essential. Numerous studies have explored voxel-level estimation techniques following image reconstruction. These methods include the random forest-based regression approach[8], the mapping-based sparse representation method[2], the semi-supervised triple dictionary learning method[9], and the multi-level canonical correlation analysis framework[10]. While these existing methods have demonstrated promising results, they tend to produce overly smoothed images.

In recent years, deep learning techniques have been extensively explored in the domain of medical imaging[11]. Generative Adversarial Networks and Convolutional Neural Networks have been proven to make a success of denoising low-dose CT images[12][13]. As deep learning is widely used in various fields, it has made a big difference in L-PET images tasks[14]. Xiang et al. proposed a sophisticated convolutional neural network model with auto-context learning that predicts high-dose PET images using only 1/4 dose of the full-dose PET images and their corresponding MR T1-images[15]. Wang et al. developed a comprehensive framework using 3D conditional Generative Adversarial Networks to generate superior-quality PET images from corresponding L-PET images[16]. Kaplan and Zhu presented a model incorporating particular image characteristics into the loss function to denoise 1/10 dose of the full-dose PET image slices and predict their full-dose counterparts[17]. Chen et al. proposed synthesizing PET images of superior quality and precision utilizing either PET-only data or a combination of PET and MR information[18]. Ouyang et al. proposed that a Generative Adversarial Network can achieve similar performance levels even in the absence of MR information[19]. More recent work by Yu et al. introduced a streamlined framework for L-PET images reconstruction, capable of rapidly producing F-PET images and leveraging the spatial details of the generated F-PET slices to improve the overall quality of the final 3D F-PET images[20].

However, there are still have several limitations, such as how we can enhance the comprehension of semantic information from different views in L-PET reconstruction, and so on. To tackle these concerns, we introduced a novel framework for L-PET reconstruction, including the diffusion generator and the u-net discriminator. We validated it by our experiment and the results showed that our approach demonstrates superiority compared to other reconstruction methods for L-PET images.

Materials and Methods

Development of GAN

A Generative Adversarial Network is widely recognized as a zero-sum game involving two network structures: a generator, which may consist of an autoencoder or its improved variations, and a discriminator, typically an autoencoder or a Convolutional Neural Network (CNN)-based discriminator. Since 2017, GAN-based models were chosen in the realm of medical image synthesis swiftly[21]. The

proliferation of GAN-based models can be attributed to their adaptable components and the increasing availability of Graphics Processing Unit (GPU) resources. In this paper, some GAN-based models have been included in the Table 1 to analyze and reproduce for the purpose of comparison. Many contemporary GAN-based models rely on the image-to-image translation technique introduced by Isola et al[22]. This method integrated the conditional GAN loss, as proposed by Mirza and Osindero[23], with an L1 regularizer loss. Consequently, the network acquires the transformation from input to target image and comprehends the loss function, enabling the generation of images closely resembling the ground truth. The conditional GAN loss is mathematically formulated as follows:

$$\mathcal{L}_{CGAN}(D, G) = -E_{x,y}[(D(x, y) - 1)^2] - E_{x,z}[D(x, G(x, z))^2]$$

1

where $z \sim p(z)$ is random noise, x is input and y is target. The L1 regularization loss can be described as followed:

$$\mathcal{L}_{L1}(G) = E_{x,y \sim P(x,y), z \sim P(z)}[||y - G(x, z)||_1]$$

2

Therefore, Considering Eqs. (1), Eqs. (2), final objective function can be written as:

$$G^*, D^* = \arg \min_G \max_D \mathcal{L}_{CGAN}(G, D) + \lambda \mathcal{L}_{L1}(G)$$

3

where the λ is a hype-parameter to equalize losses.

The CycleGAN is employed for high-resolution image-to-image translation utilizing both paired and unpaired data[24]. It comprises two generators and two discriminators. G_{AB} transfers image A to image B and G_{BA} makes the opposite transformation of G_{AB} . At the same time, D_A and D_B decide the domain of the image. The adversarial loss function for G_{AB} and D_B pair is expressed as:

$$\mathcal{L}_{GAN}(G_{AB}, D_B) = E_{b \sim P_B(b)}[\log D_B(b)] + E_{a \sim P_A(a)}[1 - \log(D_B(G_{AB}(a)))]$$

4

The adversarial loss for G_{BA} and D_A is represented as $\mathcal{L}_{GAN}(G_{BA}, D_A)$ similarly. The cyclic-consistency loss in CycleGAN is:

$$\mathcal{L}_{cyc}(G_{AB}, G_{BA}) = E_{a \sim P_A(a)}[||a - G_{BA}(G_{AB}(a))||_1] + E_{b \sim P_B(b)}[||b - G_{AB}(G_{BA}(b))||_1]$$

5

Considering the above of both, the overall loss of this model can be expressed as:

$$\mathcal{L}(G_{AB}, G_{BA}, D_A, D_B) = \mathcal{L}_{GAN}(G_{AB}, D_B) + \mathcal{L}_{GAN}(G_{BA}, D_A) + \mathcal{L}_{cyc}(G_{AB}, G_{BA})$$

6

Therefore, the objective function of the CycleGAN is:

$$G_{AB}^*, G_{BA}^* = \arg \min_{G_{AB}, G_{BA}} \max_{D_A, D_B} \mathcal{L}(G_{AB}, G_{BA}, D_A, D_B)$$

7

Architecture of DMGAN

We presented a new framework based on generative adversarial network called Diffused Multi-scale Generative Adversarial Network, including two modules: the diffusion generator and the u-net discriminator, as shown in Fig. 1. Specific structure will be introduced later. For PET images in clinical situation, our objective is to generate images that closely resemble the original F-PET images from L-PET images. The advantages of our proposed method are generating an output sample, which uses a recognized distribution of actual images and make full use of its variability. Firstly, original low-dose PET image slices are fed into the diffusion generator to synthesize the full-dose images, which also contains a sequence of corresponding target slices. In the diffusion generator, noised L-PET images are generated by introducing noise into the input L-PET images. And then, the diffusion generator generates F-PET images and noised F-PET images. The aim of this step is to enhance the generalization ability of the generator to the image and enhance the stability of the training process. Generated images are inputted into the u-net discriminator, extracting details from both overall and specific perspectives to enhance the quality of the synthesized F-PET images.

The diffusion generator

Different from the previous generative adversarial networks like a u-net generator[25] or a ResNet Generator[26][27], which were widely used in previous studies. A new generator based on the generator in transGAN[20] has been designed as shown in Fig. 2. The best advantage of this model is to get different information from different levels, enhancing the generator's generalization capability to get information from the original images. In other words, the diffusion generator provides another angle to learn original image distribution relative to the other common generators. While training DMGAN, the adversarial loss function is expressed as follows:

$$\mathcal{L}_{GAN}(D, G) = -E_{x,y}[(D(x, y) - 1)^2] - E_x[D(x, G(x))^2]$$

8

where x is representative of L-PET images, $G(x)$ is representative of F-PET images generated by the generator and y is corresponding F-PET images. We use Charbonnier Loss[28] to penalize the Euclidean disparity not only between generated F-PET images and original F-PET images:

$$\mathcal{L}(G) = E_{x,y}[\sqrt{\|y - G(x)\|^2 + \text{upvarepsilon}^2}]$$

9
but also generated noised F-PET images and original F-PET images:

$$\mathcal{L}(G_{noised}) = E_{x,y}[\sqrt{||y - G_{noised}(x)||^2 + \text{\textcolor{red}{\upvarepsilon}}^2}]$$

10
Inspired by[29][30][31][32][33][34], and considering the perceptual difference between generated F-PET images, generated noised F-PET images and original F-PET images, we adopt a VGG16-Net that is trained on the ImageNet[35] to extract feature representations of $G(x)$ and y . The perceptual loss is:

$$\mathcal{L}_{perc}(G, V) = E_{x,y}[\sqrt{||V(y) - V(G(x))||^2 + \text{\textcolor{red}{\upvarepsilon}}^2}]$$

11
where V denotes the feature maps extracted by VGG16-Net. Based on the equations above, the total loss of diffusion generator is expressed as:

$$\mathcal{L}_{dmgan}(G, V, D) = \mathcal{L}_{GAN}(D, G) + \alpha(\mathcal{L}(G) + \mathcal{L}(G_{noised})) + \beta\mathcal{L}_{perc}(G, V)$$

12
The hyper-parameters α and β are both 100.

The u-net discriminator

U-nets[36] have showcased state of art performance in numerous intricate image segmentation assignments. In these models, the encoder downsamples the input to capture the global information. And then the decoder performs upsampling. Skip connections transfer data between the encoder and decoder. Inspired by the recent GAN studies[37], we introduce a u-net discriminator in our model to get original images information from global and local views, which can cooperate the diffusion generator better. Unlike the original u-net architecture, the u-net discriminator consists of the original downsampling network and a new upsampling network, which are connected by the skip-connection and bottleneck. Contrast the original networks, the u-net discriminator performs on a per-pixel basis. The loss for the encoder is:

$$\mathcal{L}_{D_{enc}} = -E_x[\log D_{enc}(x)] - E_z[\log(1 - D_{enc}(G(z)))]$$

13
And the loss for the decoder as the mean of the all pixels can be expressed as:

$$\mathcal{L}_{D_{dec}} = -E_x[\sum_{i,j} \log [D_{dec}(x)]_{i,j}] - E_z[\sum_{i,j} \log(1 - [D_{dec}(G(z))]_{i,j})]$$

14

In Eqs. (14), $[D_{dec}(x)]_{i,j}$ and $[D_{dec}(G(z))]_{i,j}$ are represented discriminator's decision at pixel (i, j). The per-pixel outputs of D_{dec} are obtained by integrating specific details from lower-level features, enabled by skip connections from intermediate layers of the encoder, with global information derived from high-level features through upsampling from the bottleneck.

Considering Eqs. (13) and Eqs. (14), the generator objective is:

$$\mathcal{L}_{Gu-net} = -E_z[\log D_{enc}(G(z)) + \sum_{i,j} \log [D_{dec}(G(z))]_{i,j}]$$

15

encouraging the generator to concentrate on synthesizing images by capturing both global structures and local details effectively, aiming to deceive the discriminator more potent.

Inspired by[22], the loss of basic discriminator is:

$$\mathcal{L}_{adv} = E_{x,y}[\log D_{dec}(x, y)] + E_x[\log(1 - D_{dec}(G(x), x))]$$

16

In our proposed method, the u-net diffusion discriminator is return two values, representing the decoder and the encoder's output. The middle loss is used to describe the loss of the u-net discriminator encoder, which is expressed as:

$$\mathcal{L}_{middle} = -\log(1 - D_{encfake}) - \log(D_{encreal})$$

17

where $D_{encfake}$ and $D_{encreal}$ are the scalar outputs of the encoder. Therefore, the overall loss of diffusion discriminator is:

$$\mathcal{L}_d = (\mathcal{L}_{adv} + \mathcal{L}_{middle}) * 0.5\nu$$

18

where ν is a hyper-parameter set to be 1.

Datasets

The experimental dataset utilized in this study comprises 45 pediatric subjects diagnosed with epilepsy, encompassing both low-dose brain PET scans and corresponding full-dose brain PET scans, which are collected in 2020. The FDG-PET images of all subjects' brains are obtained using the whole-body hybrid PET/MR system (SIGNA PET/MR, GE Healthcare). In the context of clinical practice, we did not exclude images exhibiting comparatively lower quality. L-PET images are generated by reconstructing list-mode F-PET data, which undergoes a 5% undersampling process. Using 5% dose can significantly reduce radiation risks, making the research results more valuable in actual clinical practice. And the 5% dose

selection provides an extremely challenging testing environment to verify the ability of our reconstruction algorithm to generate high-quality images under extremely low-dose conditions. This rigorous testing helps demonstrate the robustness and effectiveness of the algorithm. Prior to reconstruction, both L-PET and F-PET images are subjected to preprocessing using Statistical Parametric Mapping for realignment and normalization. Following this preprocessing, the voxel size of these PET images is standardized to $1 \times 1 \times 1$ mm. This voxel size is chosen to provide higher spatial resolution. This is particularly important for detecting and analyzing fine structures, especially during the reconstruction of low-dose PET images. High-resolution images help improve image quality and the presentation of details. Through the processing of three-dimensional brain images within the PET dataset, we derived 256×256 2D brain image slices, which served as the basis for our experimental analyses.

Data analysis

In this study, to illustrate the effectiveness of the DMGAN for L-PET images reconstruction, we calculated the metrics of the model on the dataset, including SSIM and PSNR. SSIM is used to assess the similarity between the reconstructed image and the reference image. We used images of the entire brain region to calculate SSIM. This is because the structural characteristics of the entire brain region are important for evaluating image quality. The calculation of the SSIM value is based on the local window of each voxel, and the similarity of the image is evaluated by comparing the brightness, contrast and structural information. PSNR is used to assess the noise level between the reconstructed image and the reference image. We also used the image of the entire brain region to calculate PSNR. This is because the evaluation of the noise level needs to consider the global characteristics of the entire image. Meanwhile, the reconstruction ability of each model to reconstruct L-PET images was analyzed in the same way. The analysis and comparison of pseudo-color difference images were carried out through OpenCV package to further intuitively show the difference in the results of each method. To mitigate the influence of diverse initialization parameters on experimental outcomes, we set the same random seeds for all experimental methods. We utilized the PyTorch library to conduct this experiment.

Results

Experimental setting

The experimental dataset comprised 45 pediatric patients diagnosed with epilepsy, encompassing both brain L-PET scans and corresponding brain F-PET scans. By processing the three-dimensional brain images in the PET dataset, we can obtain 256×256 2D brain image slices, which can be used in our experiments. The dataset was divided randomly into training, validation, and testing sets, with 80% for training, 10% for validation, and 10% for testing, to ensure diversity and representativeness. To validate the effectiveness of the proposed method, we compared it with cGAN[23], CycleGAN[24], and transGAN[20]. And in order to eliminate the impact of different initialization parameters on experiments, we set the same random seeds for all experimental methods. We used Pytorch library on a NVIDIA RTX 4090 GPU. The batch size was 1 and the epoch was set to be 300. Finally, we evaluated them through

qualitative evaluation and quantitative measures. For quantitative comparisons, we chose two metrics to evaluate performance of various methods: SSIM and PSNR.

Quantitative Evaluation

As shown in Table 1, it is evident that the proposed method surpasses all other compared approaches, attaining the highest PSNR and SSIM values. These results suggest that the images generated by our method closely resembles that of corresponding original F-PET images. It improves PSNR by at least 6.2% over the other methods, which demonstrates its superiority.

Table 1
Quantitative compared results of DMGAN and other methods for L-PET reconstruction.

Method	L-PET	cGAN	CycleGAN	transGAN	DMGAN
PSNR	23.02	27.28 ± 0.42	27.08 ± 0.34	31.90 ± 0.16	33.89 ± 0.14
SSIM	0.731	0.831 ± 0.007	0.829 ± 0.008	0.897 ± 0.002	0.911 ± 0.002

Qualitative Evaluation

In Fig. 3, some random samples are selected to compare. We can observe that our proposed method has better accurate structural details than other approaches. And compared with F-PET images, there are more metabolic details in our approach than other ways. While SSIM and PSNR are standard measures to quantify image quality in computer vision, in nuclear medicine, a physician's observer score is often more relevant and necessary. Deep learning methods can sometimes introduce structural changes that do not significantly affect these quantitative measures but may still be important for clinical diagnosis. Therefore, a more in-depth evaluation of the images, such as expert assessments, is valuable. To address this concern, we included evaluations from an experienced nuclear medicine physician. The opinion scores aim to evaluate clinical feasibility, specifically concerning metabolic details of the brain in generated F-PET images. Test set images are provided to the physician for evaluation in a randomized order, with all labels removed. Each PET image is rated for image quality on a five-point scale by the physician. As depicted in Fig. 4, the opinion scores for DMGAN are notably superior to those of other compared methods. It is evident that DMGAN achieves state-of-the-art results in L-PET image reconstruction and garners favorable clinical assessments. The more intuitive comparison is displayed in Fig. 5. Generated F-PET images and original F-PET images are computed in pseudo-color difference maps, revealing that the proposed method exhibited the smallest voxel-scale difference compared to the other methods. Considering the clinical significance of accurate low-dose reconstruction of the epilepsy focus in pediatric patients, Fig. 6 presents the L-PET image of the epilepsy focus, the synthesized F-PET images, and the ground truth F-PET image of the epilepsy focus. It is evident that compared to other methods, the synthesized F-PET image generated by DMGAN exhibits a more precise voxel-level metabolic intensity distribution, leading to a sharper and more detailed depiction of the epileptic focus.

Ablation study

To certificate the effectiveness of every module, we designed the ablation study by removing key modules from the final method on PET dataset in Table 2. In this part, we proved that our proposed method consists of two modules, which both make a contribution to the final result. As we can see, compared to the final method, the reduction of Diffusion Generator and u-net discriminator both decrease the metric value, respectively declines over four hundredths. This is enough to prove the effectiveness of each module of this method. Figure 7 shows the comparative results of this ablation study.

Table 2
Performance of every module in our proposed method DMGAN.

Module	PSNR	SSIM
transGAN	31.90 ± 0.16	0.897 ± 0.002
transGAN with Diffusion generator	32.48 ± 0.14	0.901 ± 0.002
transGAN with u-net discriminator	32.40 ± 0.34	0.902 ± 0.003
transGAN with Diffusion generator and u-net discriminator (DMGAN)	33.89 ± 0.14	0.911 ± 0.002

Discussion

PET is a widely utilized medical imaging technology for diagnosing various diseases. While full-dose imaging ensures high-quality images, concerns about radiation exposure persist. Balancing the need to reduce radiation exposure with maintaining diagnostic accuracy is crucial. This challenge is effectively addressed by reconstructing L-PET images to match the quality of F-PET images counterparts. In this study, we suggested DMGAN for efficient full-dose reconstruction of L-PET images. DMGAN comprises two modules: the diffusion generator and the u-net discriminator. Sequential L-PET slices is processed by the diffusion generator, generating higher quality F-PET images. The u-net discriminator learns to distinguish from global and local views between the real data, generated data and generated diffused data. Experimental results demonstrated that our proposed DMGAN achieves excellence on commonly utilized assessment criteria. We can observe the performance of the model as followed. Compared with L-PET images, it achieves 38.6% improvement in terms of PSNR and 24.6% improvement in terms of SSIM, which can prove it does have the ability to reconstruct F-PET images from L-PET images. To further demonstrate its superiority over existing models, we compared it with other methods. For example, compared with cGAN, CycleGAN and transGAN, it achieves about 24.2%, 25.1% and 6.2% improvement respectively in terms of PSNR and 9.6%, 9.9% and 1.6% improvement respectively in terms of SSIM. Our results showed that the proposed model can reconstruct L-PET images to F-PET images, thereby effectively addressing the challenge of balancing reduced radiation exposure with preserved diagnostic performance.

In previous research, the training process involves fusing conditional information into the generator and discriminator for cGAN[23]. For example, the input to both the generator and discriminator encompasses the conditional information derived from PET images, guiding the generation process and ensuring that the generated data exhibits heightened realism and controllability under the given PET images. Overall, cGAN serves as an extension to Generative Adversarial Networks, introducing PET image data information to enable the generation model to produce data tailored to the distribution of F-PET images. CycleGAN was originally designed for image translation tasks with the primary goal of learning mappings across two distinct domains without requiring paired training data[24]. Its most common application involves translating images between two domains, such as we can use it mapping L-PET to F-PET. In comparison to earlier GAN architectures, CycleGAN introduced the concept of cycle consistency loss as introduced before to address the training challenges faced by traditional GANs when dealing with unpaired data. The cycle consistency loss ensures that the two-directional image transformations performed by the generator are reversible, meaning that the transformed images can be accurately reverted back to their original form. This helps prevent the generator from producing unnatural outputs. The overall objective of training the model is to minimize the discrepancies between the PET images generated by the network and the corresponding F-PET images. TransGAN is a variant of GANs built upon the Transformer architecture, originally designed for image generation tasks[20]. Leveraging the self-attention mechanism inherent to Transformers, TransGAN aims to capture global information and dependencies within PET images. The model utilizes the multi-head attention mechanism to simultaneously focus on different regions of PET images. Both the generator and discriminator of TransGAN are constructed based on the Transformer architecture. In the context of L-PET reconstruct tasks, TransGAN exhibits greater flexibility and potential for global modeling, allowing it to effectively handle PET images.

We have deduced and explained proposed method principle. Getting different information from different views in the same images aims to improve the ability of the generator. A u-net discriminator is introduced to extract features from various views. By using u-net discriminator, we enable each layer to make an assessment so that different layers capture distinct semantic information, providing judgments on the image from a multi-scale perspective. Both of them are the main components of the structure and they enhance the quality of synthetic PET images with modified loss function. The comprehensive experiments, incorporating qualitative assessments and quantitative metrics, have substantiated that the proposed DMGAN is capable of synthesizing realistic PET images. It can restore original details from L-PET images, surpassing the performance of other models trained on identical datasets. The synthesized F-PET image generated by DMGAN can make a sharper and clearer depiction of the epilepsy focus. We also conducted ablation experiments to further prove that each module plays a certain role in the final result.

There are some limitations in our study. Firstly, to enhance performance, we have incorporated additional efficient modules, potentially necessitating increased hardware resources during runtime. Secondly, the existing methodology exclusively focuses on PET images. Future efforts can involve the integration of medical images from other modalities to further optimize outcomes.

Conclusion

The study introduced an innovative framework based on generative adversarial networks for reconstructing low-dose PET images. At the meantime, we proposed a diffusion generator which can generate higher quality reconstructed images, which can be more approached to full-dose PET images. The final experimental results demonstrated that our method has the superiority in quantitative and in qualitative evaluation compared with other reconstruction methods of low-dose PET images. This study can help to mitigate patients' radiation exposure while maintaining clinical diagnostic efficacy.

Declarations

Ethics approval and consent to participate

This study was performed in line with the principles of the Declaration of Helsinki. This retrospective study received approval from the institutional review board of Hangzhou Universal Medical Imaging Diagnostic Center and the requirement of informed consent was waived (Approval No. [2021] 001).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability statement

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Funding

This study was supported by the National Natural Science Foundation of China (82030049), the National Key Research and Development Program of China (2021YFA1101700, 2021YFE0108300) and the Fundamental Research Funds for the Central Universities.

Author contributions

XY contributed to designing the study, acquiring the data, developing the deep learning algorithm, analysing the study data and drafting the manuscript. DYH helped analysing the data and revised the manuscript. QY helped revising the manuscript. YF contributed to the study design and data acquisition. YZ and JW contributed to the study design. MT and HZ contributed to the study design, data acquisition and revising the manuscript. All authors read and approved the final manuscript.

Acknowledgements

None.

References

1. Tian M, He X, Jin C, et al. Transpathology: molecular imaging-based pathology[J]. European journal of nuclear medicine and molecular imaging, 2021, 48: 2338-2350.
2. Wang Y, Zhang P, An L, et al. Predicting standard-dose PET image from low-dose PET and multimodal MR images using mapping-based sparse representation[J]. Physics in Medicine & Biology, 2016, 61(2): 791.
3. Tan H, Sui X, Yin H, et al. Total-body PET/CT using half-dose FDG and compared with conventional PET/CT using full-dose FDG in lung cancer[J]. European journal of nuclear medicine and molecular imaging, 2021, 48: 1966-1975.
4. Boellaard R, Delgado-Bolton R, Oyen W J G, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0[J]. European journal of nuclear medicine and molecular imaging, 2015, 42: 328-354.
5. Zhou L, Schaefferkoetter J D, Tham I W K, et al. Supervised learning with cyclegan for low-dose FDG PET image denoising[J]. Medical image analysis, 2020, 65: 101770.
6. Wang Y R, Baratto L, Hawk K E, et al. Artificial intelligence enables whole-body positron emission tomography scans with minimal radiation exposure[J]. European journal of nuclear medicine and molecular imaging, 2021, 48: 2771-2781.
7. Wang C, Hu Z, Shi P, et al. Low dose PET reconstruction with total variation regularization[C]//2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2014: 1917-1920.
8. Kang J, Gao Y, Shi F, et al. Prediction of standard-dose brain PET image by using MRI and low-dose brain [18F] FDG PET images[J]. Medical physics, 2015, 42(9): 5301-5309.
9. Wang Y, Ma G, An L, et al. Semisupervised triple dictionary learning for standard-dose PET image prediction using low-dose PET and multimodal MRI[J]. IEEE transactions on biomedical engineering, 2016, 64(3): 569-579.
10. An L, Zhang P, Adeli E, et al. Multi-level canonical correlation analysis for standard-dose PET image estimation[J]. IEEE Transactions on Image Processing, 2016, 25(7): 3303-3315.
11. Kawahara J, Brown C J, Miller S P, et al. BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment[J]. NeuroImage, 2017, 146: 1038-1049.
12. Chen H, Zhang Y, Kalra M K, et al. Low-dose CT with a residual encoder-decoder convolutional neural network[J]. IEEE transactions on medical imaging, 2017, 36(12): 2524-2535.
13. Yang Q, Yan P, Zhang Y, et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss[J]. IEEE transactions on medical imaging, 2018, 37(6): 1348-1357.
14. Shan H, Zhang Y, Yang Q, et al. 3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network[J]. IEEE transactions on medical imaging, 2018, 37(6): 1522-1534.

15. Xiang L, Qiao Y, Nie D, et al. Deep auto-context convolutional neural networks for standard-dose PET image estimation from low-dose PET/MRI[J]. *Neurocomputing*, 2017, 267: 406-416.
16. Wang Y, Yu B, Wang L, et al. 3D conditional generative adversarial networks for high-quality PET image estimation at low dose[J]. *Neuroimage*, 2018, 174: 550-562.
17. Kaplan S, Zhu Y M. Full-dose PET image estimation from low-dose PET image using deep learning: a pilot study[J]. *Journal of digital imaging*, 2019, 32(5): 773-778.
18. Chen K T, Gong E, de Carvalho Macruz F B, et al. Ultra-low-dose 18F-florbetaben amyloid PET imaging using deep learning with multi-contrast MRI inputs[J]. *Radiology*, 2019, 290(3): 649-656.
19. Ouyang J, Chen K T, Gong E, et al. Ultra-low-dose PET reconstruction using generative adversarial network with feature matching and task-specific perceptual loss[J]. *Medical physics*, 2019, 46(8): 3555-3564.
20. Fu Y, Dong S, Liao Y, et al. A resource-efficient deep learning framework for low-dose brain PET image reconstruction and analysis[C]//2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI). IEEE, 2022: 1-5.
21. Wang T, Lei Y, Fu Y, et al. A review on medical imaging synthesis using deep learning and its clinical applications[J]. *Journal of applied clinical medical physics*, 2021, 22(1): 11-36.
22. Isola P, Zhu J Y, Zhou T, et al. Image-to-image translation with conditional adversarial networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1125-1134.
23. Mirza M, Osindero S. Conditional generative adversarial nets[J]. *arXiv preprint arXiv:1411.1784*, 2014.
24. Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2223-2232.
25. Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros, "Image-to-image translation with conditional adversarial net- works," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1125–1134.
26. Dar S U H, Yurt M, Karacan L, et al. Image synthesis in multi-contrast MRI with conditional generative adversarial networks[J]. *IEEE transactions on medical imaging*, 2019, 38(10): 2375-2388.
27. Hu S, Shen Y, Wang S, et al. Brain MR to PET synthesis via bidirectional generative adversarial network[C]//Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part II 23. Springer International Publishing, 2020: 698-707.
28. Jonathan T Barron, "A general and adaptive robust loss function," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 4331–4339.
29. Jiahong Ouyang, Kevin T Chen, Enhao Gong, John Pauly, and Greg Zaharchuk, "Ultra-low-dose pet reconstruction using generative adversarial network with feature matching and task-specific perceptual loss," *Medical physics*, vol. 46, no. 8, pp. 3555–3564, 2019.

30. Yang Lei, Xue Dong, Tonghe Wang, et al. "Whole-body pet estimation from low count statistics using cycle-consistent generative adversarial networks," *Physics in Medicine & Biology*, vol. 64, no. 21, pp. 215017, 2019.
31. Karl Spuhler, Mario Serrano-Sosa, Renee Cattell, Christine DeLorenzo, and Chuan Huang, "Full-count pet recovery from low-count image using a dilated convolutional neural network," *Medical Physics*, vol. 47, no. 10, pp. 4928–4938, 2020.
32. Karen Simonyan and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
33. Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.
34. Salman UH Dar, Mahmut Yurt, Levent Karacan, Aykut Erdem, Erkut Erdem, and Tolga Cukur, "Image synthesis in multicontrast mri with conditional generative adversarial networks," *IEEE transactions on medical imaging*, vol. 38, no. 10, pp. 2375–2388, 2019.
35. Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition. Ieee*, 2009, pp. 248–255.
36. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//*Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III* 18. Springer International Publishing, 2015: 234-241.
37. Schonfeld E, Schiele B, Khoreva A. A u-net based discriminator for generative adversarial networks[C]//*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020: 8207-8216.

Figures

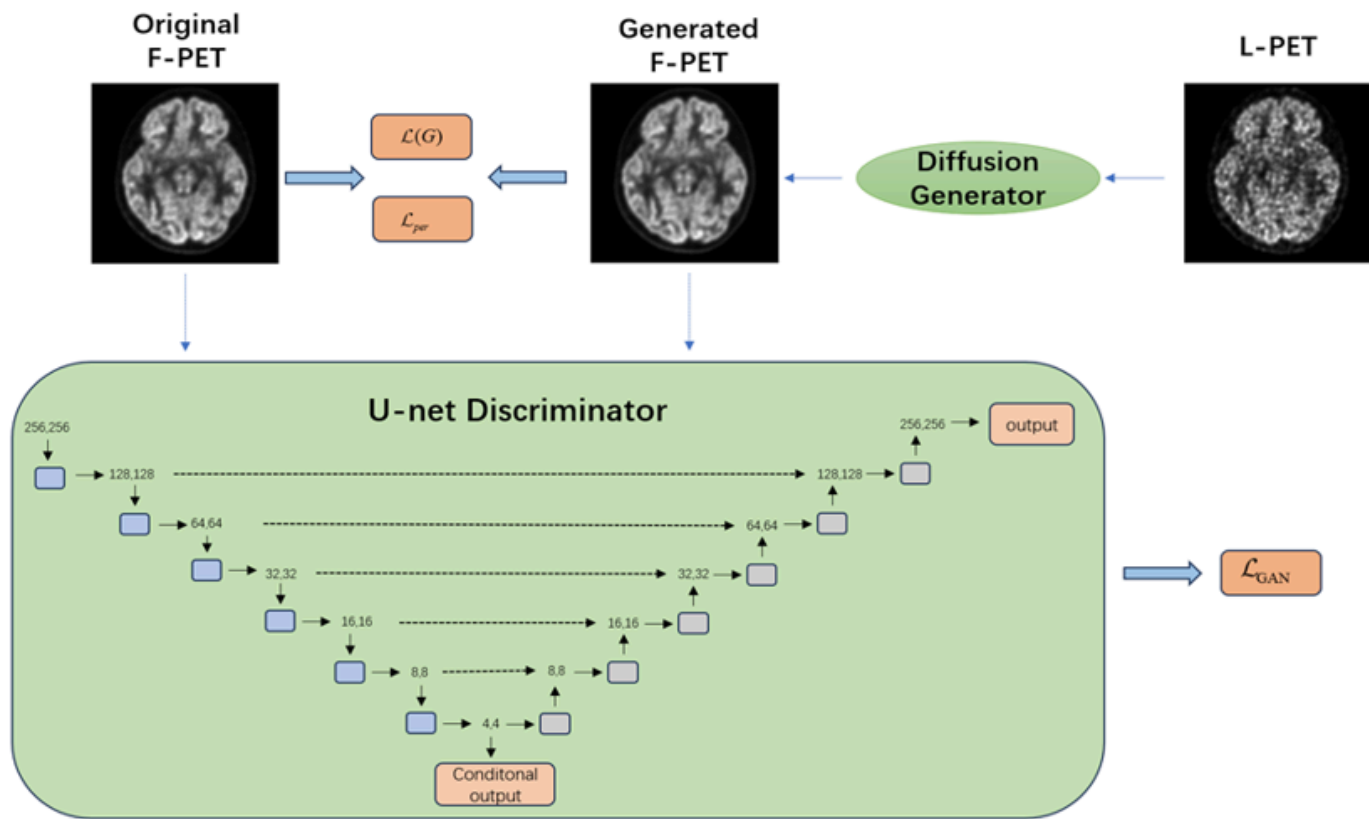


Figure 1

An overview of the proposed DMGAN for L-PET images reconstruction.

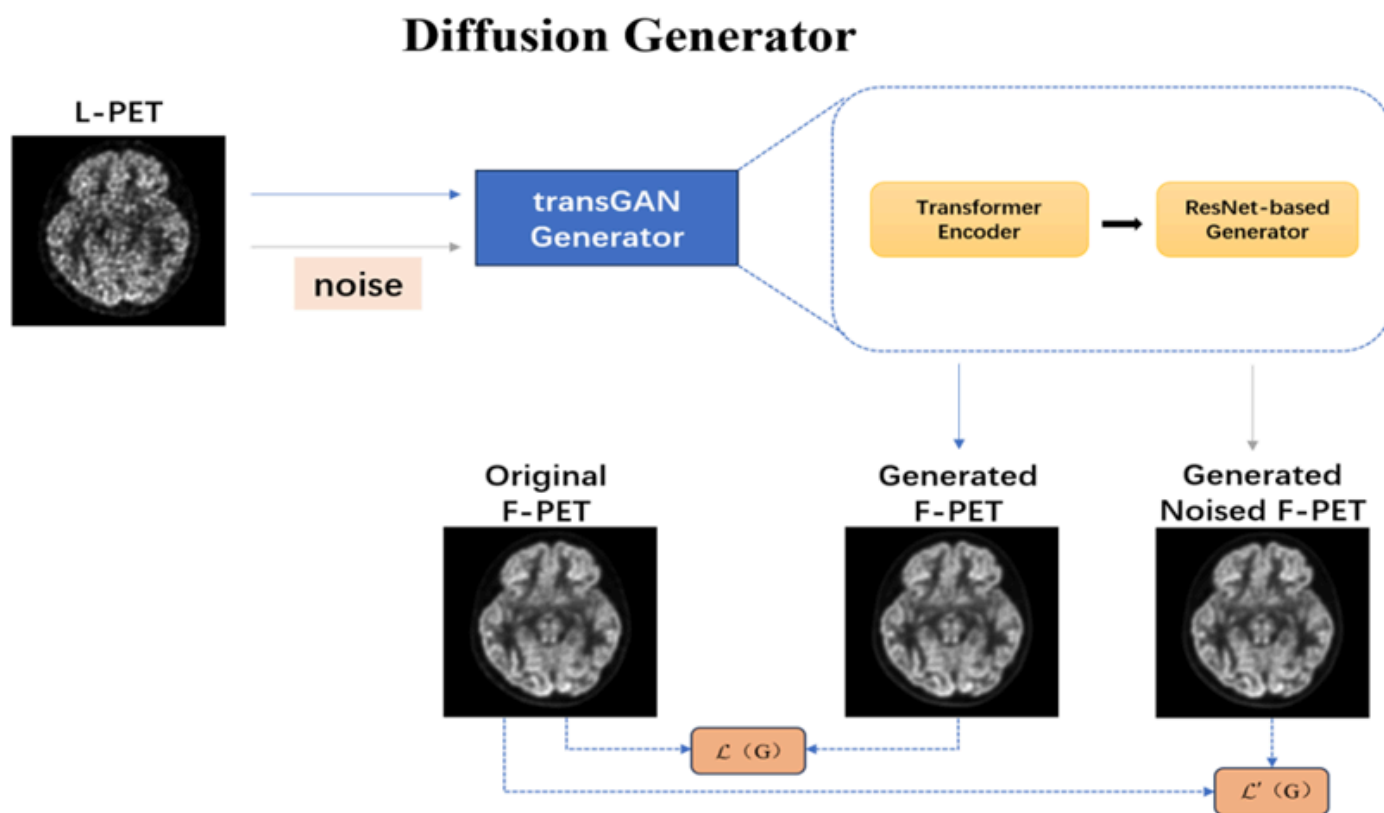


Figure 2

The architecture of the proposed diffusion generator for generated images.

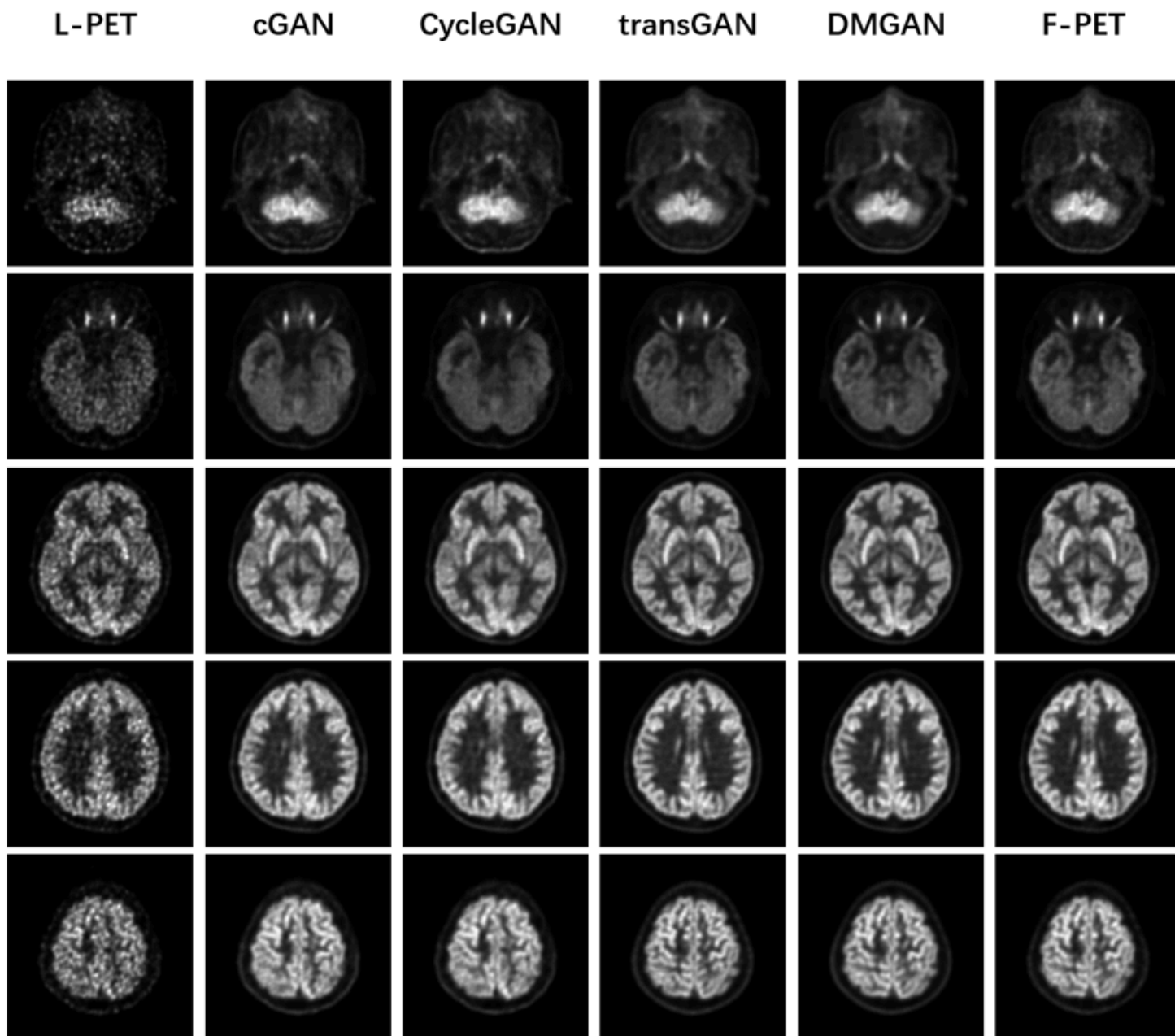


Figure 3

Qualitative evaluations of various methods on L-PET images.

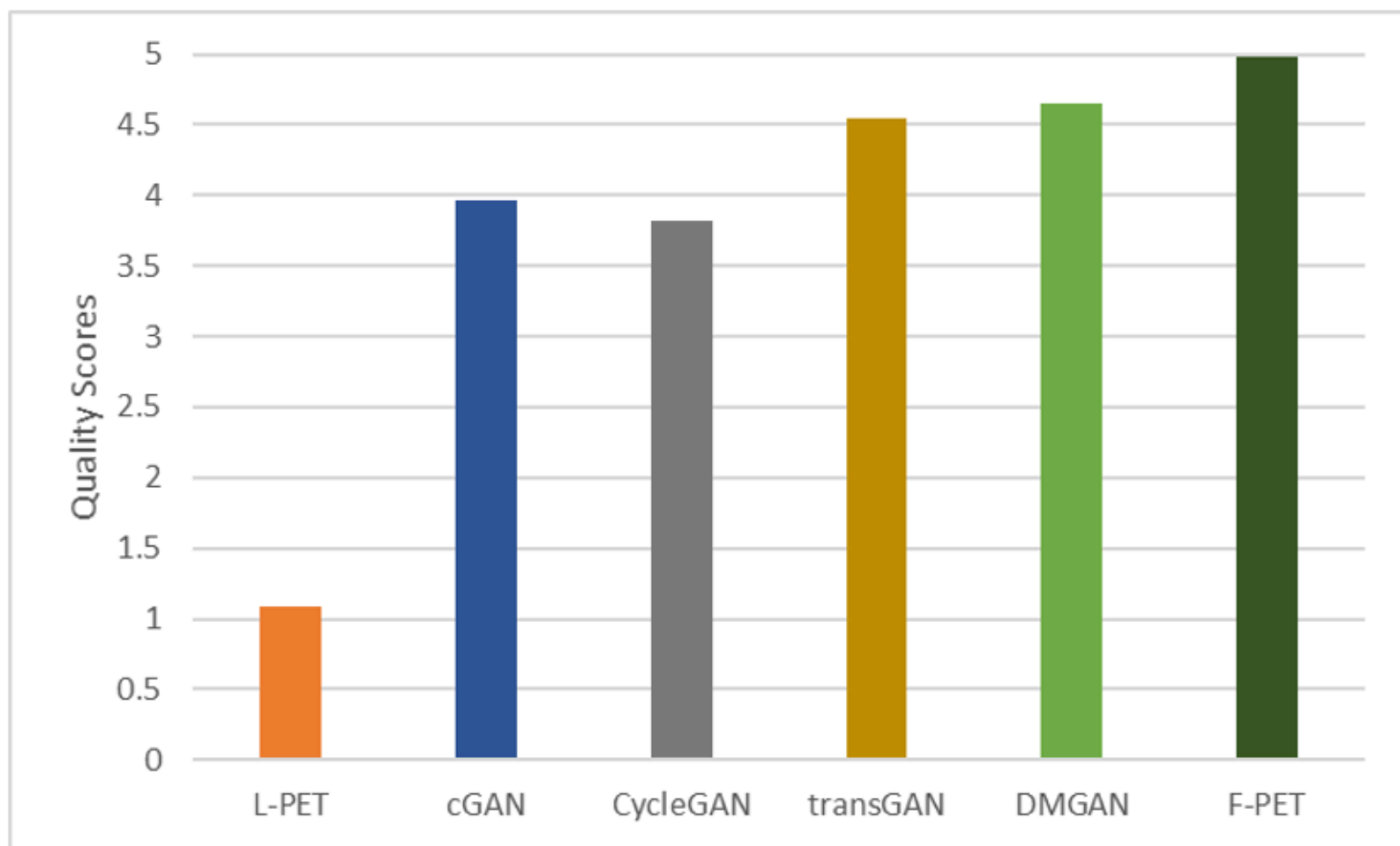


Figure 4

Physician performed clinical evaluations of synthetic F-PET images generated by our method and the comparative models

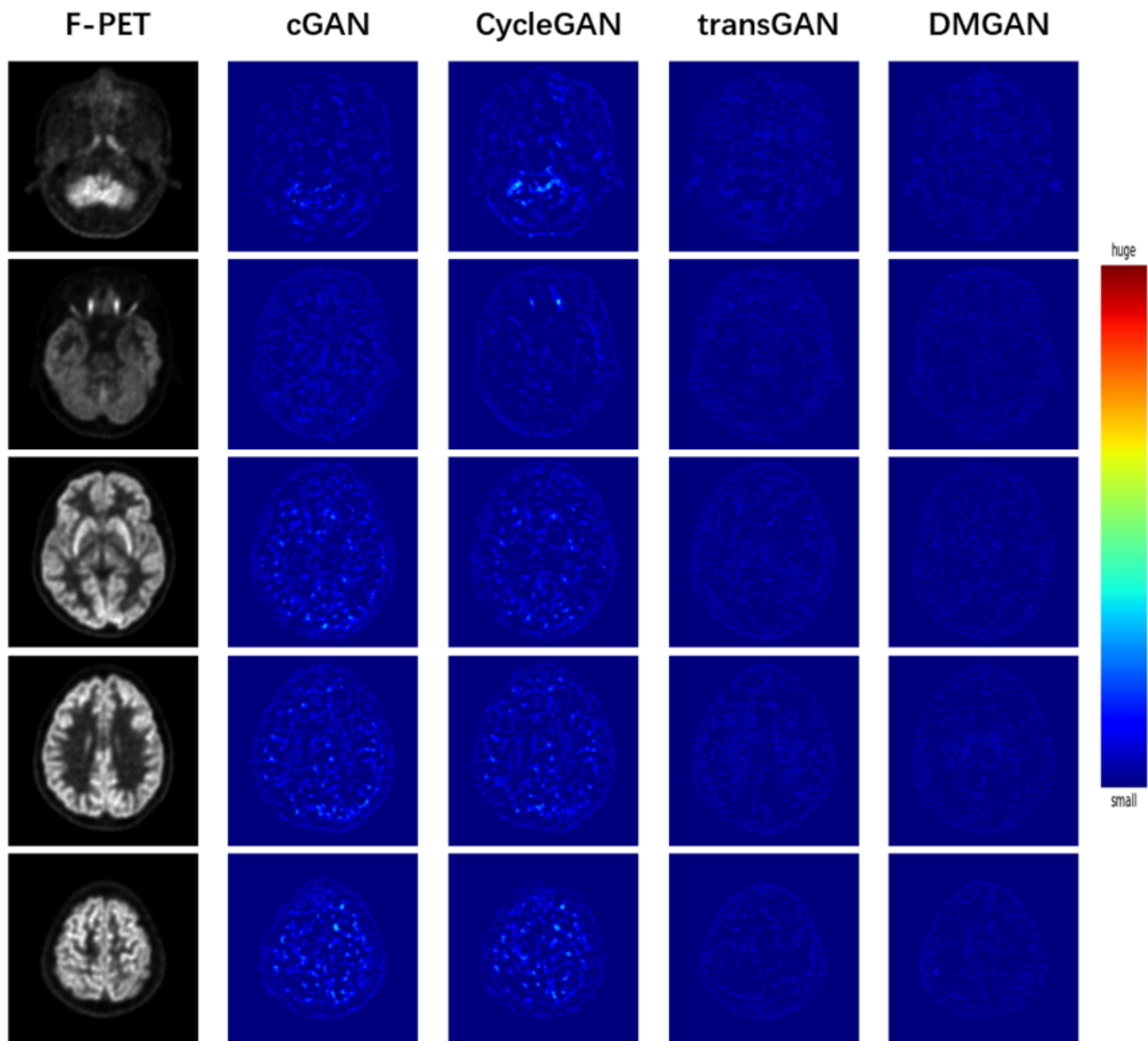


Figure 5

Comparison of the pseudo-color variance maps among the generated images and the original F-PET images from existing methods and our method. The colors transition from blue to red, indicating absolute differences ranging from small to huge.

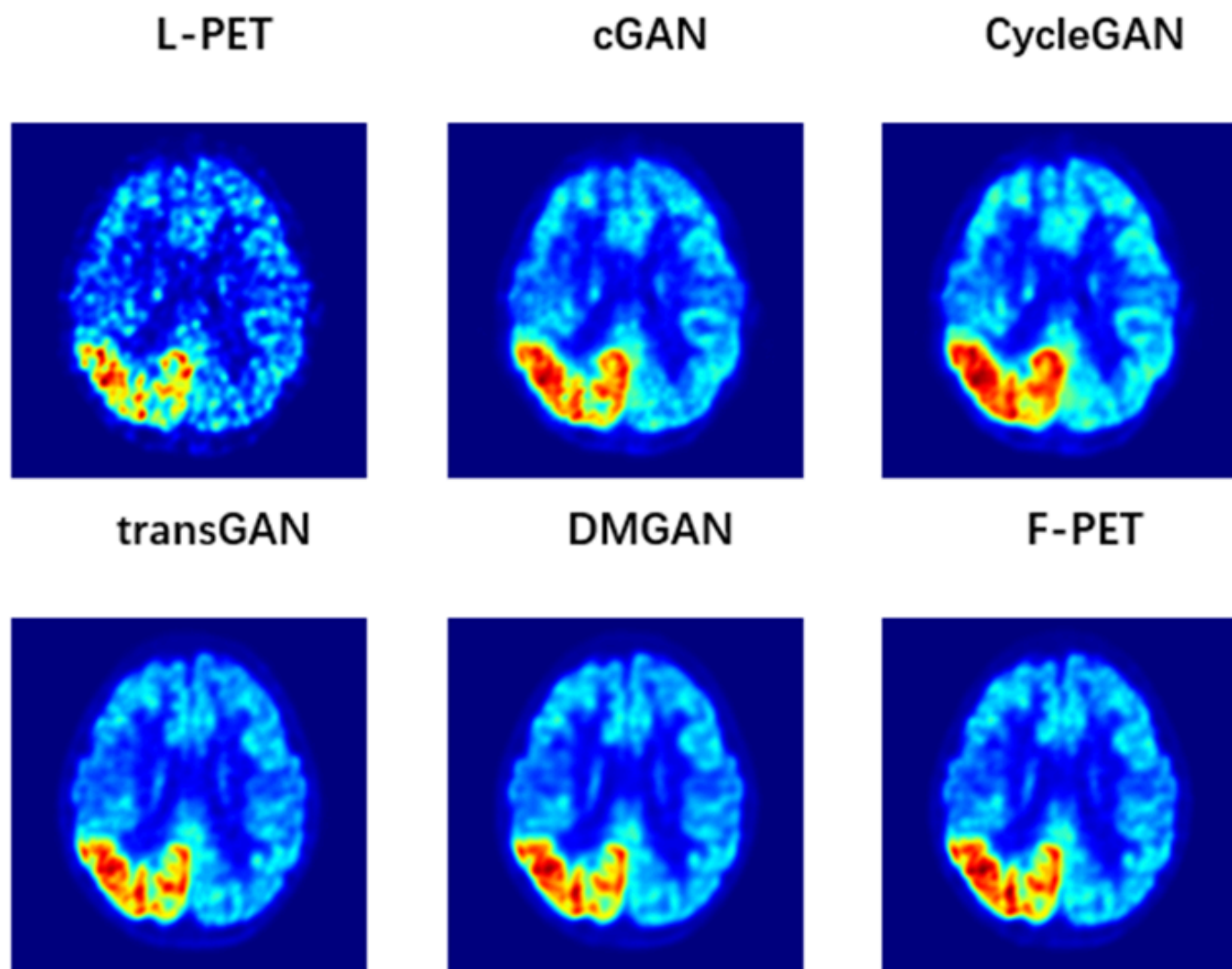


Figure 6

Displaying the epilepsy focus L-PET images, the synthesized epilepsy focus F-PET images generated by various methods, and the ground truth epilepsy focus F-PET images. The colour gradient ranges from blue to red, representing metabolic values of FDG PET images from low to high.

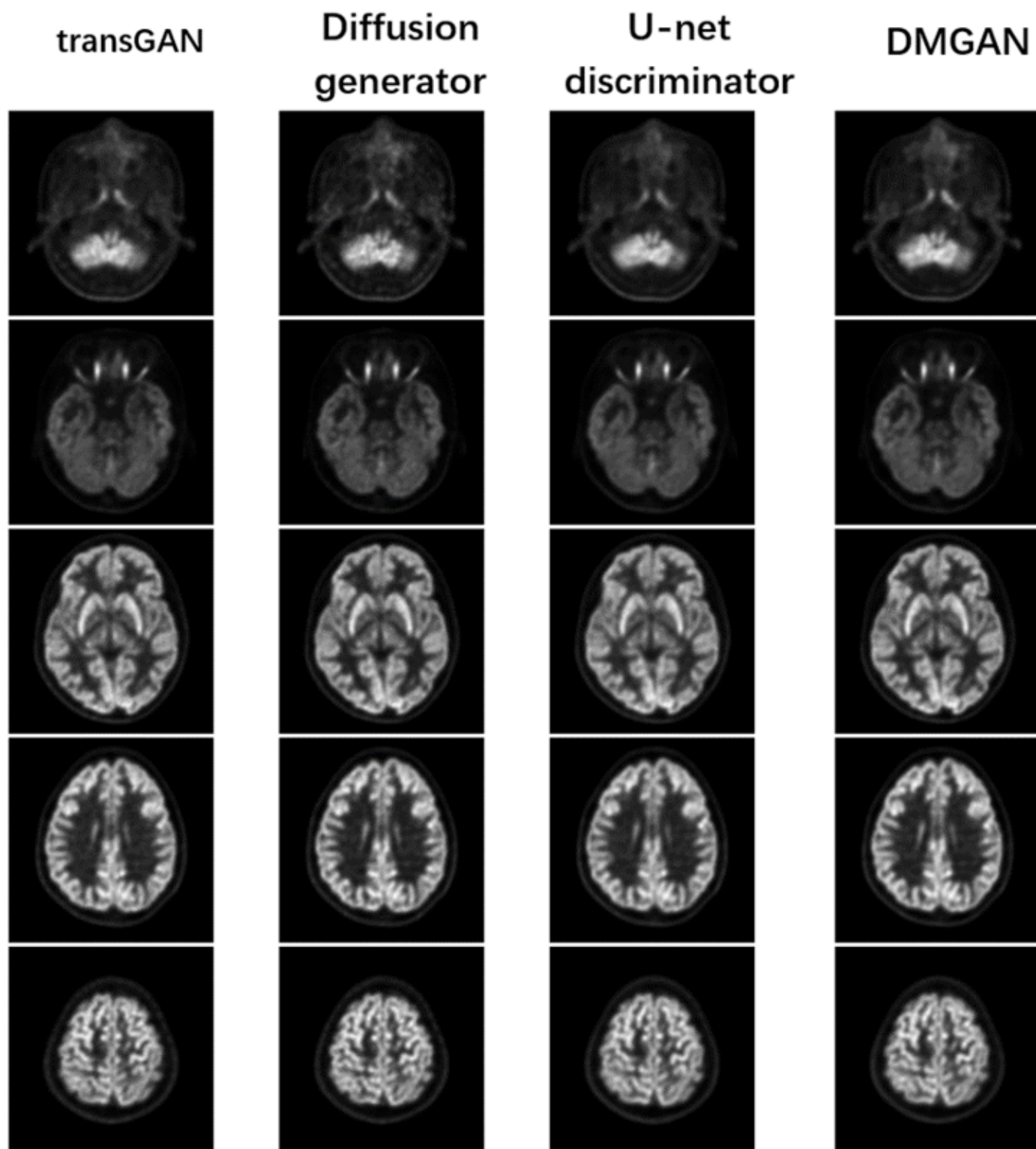


Figure 7

Comparative results from the ablation studies.