RegC2P: A registration-enhanced GAN for 3D CT-to-PET translation

Nguyen Thanh Trung*, Le Ke Hung

Medical Equipment Department, 108 Military Central Hospital, 1B Tran Hung Dao, Hai Ba Trung, Hanoi, Vietnam.

*Corresponding author: thanhtrungys@yahoo.com

Received 24 Dec. 2024; Revised 09 Feb. 2025; Accepted 04 Apr. 2025; Published 15 Apr. 2025.

DOI: https://doi.org/10.54939/1859-1043.j.mst.102.2025.125-132

ABSTRACT

The fusion of Positron Emission Tomography (PET) and Computed Tomography (CT) has significantly advanced cancer imaging by combining metabolic and anatomical information, enhancing diagnosis, staging, and treatment monitoring. However, the widespread use of PET/CT systems is limited by the scarcity of PET scanners and the reliance on radioactive tracers. To address this, 3D image-to-image translation has emerged as a promising solution for generating synthetic PET images from CT scans. Existing generative methods based on Generative Adversarial Networks (GANs) face challenges such as instability and stochastic outputs that lack precision for reliable 3D CT-to-PET translation. We propose RegC2P, a novel approach that integrates a registration module into a GAN-based framework to generate accurately aligned 3D PET images from CT scans. RegC2P transforms the problem into a slice-by-slice 2D image-to-image translation task, where individual 2D CT slices are translated into PET slices and then stacked into a 3D PET volume. Misaligned 2D PET slices are treated as noisy labels, and the generator is trained with an additional registration network to adaptively correct the misalignment, optimizing both the translation and registration tasks simultaneously. To ensure smoothness and consistency across generated PET slices throughout the entire volume, we introduce a 3D U-Net refinement network. Extensive experiments on large datasets demonstrate that RegC2P outperforms state-of-the-art methods, achieving a 10.16% reduction in MAE, a 0.96% improvement in SSIM, and a 3.6% increase in PSNR, setting a new benchmark for the quality of synthesized 3D PET images.

Keywords: 3D Image-to-Image translation; 3D PET/CT imaging; CT-to-PET; GAN; U-Net.

1. INTRODUCTION

The combination of PET and CT imaging has transformed cancer diagnosis and treatment strategies. PET scans reveal the metabolic and functional characteristics of tissues, whereas CT scans deliver detailed anatomical imaging. Together, these techniques provide a synergistic approach, allowing clinicians to precisely locate tumors, assess cancer stages, and evaluate the effectiveness of treatments. This synergy enhances diagnostic confidence and enables more targeted therapeutic strategies, ultimately improving patient outcomes in oncology [1]. Despite these benefits, widespread adoption of PET/CT systems is constrained by factors such as the need for radioactive materials, high operational costs, and the limited availability of PET scanners compared to the more common CT scanners [2]. Consequently, image synthesis techniques offer a promising solution by generating images for missing modalities, facilitating comprehensive multimodal image fusion analyses.

The CT-to-PET (CT2PET) translation task is a specific challenge within the broader domain of Image-to-Image (I2I) translation. Generative Adversarial Networks (GANs) [3], have emerged as the dominant approach for this task. Pix2Pix [4] introduced a generator-discriminator framework that minimizes reconstruction and adversarial losses, paving the way for models like CycleGAN [5] and SelectionGAN [6]. Despite their ability to generate synthetic images that closely resemble real ones, GANs often need to be more stable and highly sensitive to training hyperparameters, which limits their reliability. To address these limitations, many studies have been directed toward adapting I2I translation techniques to medical imaging challenges, as demonstrated by recent studies [7-10]. Despite significant advancements, several challenges persist: (1) much of the existing research has concentrated on translating between similar modalities, such as MRI T1-to-MRI T2 or CBCT-to-CT, while translating between fundamentally different modalities, such as CT2PET, remains a complex and unresolved problem; (2) while generative model-based approaches have shown promise in generating images consistent with the target domain's distribution, their reliance on source image guidance often falls short due to intrinsic randomness in outputs, resulting in varying image quality.

In this study, we address the critical yet often underexplored task of CT2PET translation, aimed at improving diagnostic accuracy in oncology. Through a comprehensive review of the existing literature, we identify two primary challenges in CT2PET translation: (a) achieving accurate anatomical alignment between the two modalities, and (b) ensuring consistency across modalities to effectively capture relevant structural information. To address these challenges, we propose a novel method, referred to as Registration-enhanced CT-to-PET (RegC2P). The central concept of RegC2P involves reducing the 3D CT2PET translation problem to a slice-by-slice 2D CT2PET translation task. The individual 2D PET slices generated are subsequently stacked to form a 3D volume. In RegC2P, misaligned PET slices are treated as noisy labels, framing the problem as supervised learning with noisy labels. Additionally, to ensure smoothness and consistency across the entire volume, we integrate a 3D U-Net refinement network to refine the generated slices. By integrating these targeted strategies, RegC2P effectively mitigates the limitations highlighted in prior studies, providing a robust and precise framework for advancing 3D CT2PET translation.

The main contributions of this paper include:

• We introduce RegC2P, a GAN-based model that incorporates a registration technique to address the challenging task of synthesizing misaligned PET images from CT inputs. This task is particularly difficult due to the inherent differences in anatomical structure and functional information between the CT and PET modalities. To the best of our knowledge, RegC2P is the first model to utilize GANs to tackle these cross-modality challenges in the context of 3D imaging.

• To enhance the quality and diagnostic relevance of the generated PET images, we introduce a novel use of a registration regularization component that aligns PET outputs with CT inputs, reducing spatial misalignment and enhancing structural coherence across modalities. Furthermore, we incorporate a 3D U-Net refinement network to ensure smooth transitions and consistency between adjacent PET slices, maintaining anatomical integrity throughout the entire volume. This combination of registration-based alignment and refinement enhances both the accuracy and the clinical applicability of the synthesized PET images, making RegC2P a promising tool for improving diagnostic precision in oncology.

• We introduce the 3D PET/CT dataset, which comprise 1,000 paired 3D CT-PET images. The dataset is carefully curated to facilitate the development and evaluation of advanced medical image translation techniques, offering a valuable resource for the research community.

• We conduct comprehensive experiments to evaluate the performance of RegC2P against state-of-the-art image translation methods, using the proposed dataset. The results show that RegC2P surpasses previous methods in PET image quality, achieving a 15.5% lower MAE, 1.2% higher SSIM, and 0.4% higher PSNR compared to the best-performing benchmark models across both datasets.

2. RELATED WORK

I2I translation has been widely studied in medical imaging, particularly for generating synthetic images between different modalities. GANs have dominated the field, with two primary approaches:

supervised and unsupervised methods. Supervised methods such as Pix2Pix [4] require paired and well-aligned datasets, which are often challenging to obtain due to respiratory motion or anatomical changes between scans. Unsupervised approaches like CycleGAN [5] overcome this limitation by learning from unpaired data but often fail to achieve optimal performance. RegGAN [10] was introduced to address the limitations of these approaches by treating misaligned target images as noisy labels. To achieve this, RegGAN integrates a registration network into the generator, enabling it to adaptively correct the misalignment during training. By jointly optimizing the translation and registration tasks, RegGAN achieves robust performance on both aligned and misaligned datasets. However, RegGAN operates exclusively on 2D images, which limits its ability to ensure spatial consistency when extended to 3D medical imaging.

More recently, diffusion models have emerged as state-of-the-art generative techniques, offering superior performance for medical image generation tasks. CPDM [11] is one of the first works to utilize diffusion models for CT2PET translation. CPDM introduces two conditional maps-an Attention Map and an Attenuation Map-to guide the diffusion process, focusing on regions of clinical importance and improving diagnostic accuracy. While CPDM demonstrates impressive results in generating 2D PET images, its extension to 3D remains computationally challenging and has not been addressed in their work. To overcome the inherent inconsistencies between slices in 3D imaging, DDPET-3D [12] introduces a dose-aware diffusion model for 3D low-dose PET imaging. Unlike earlier methods, DDPET-3D explicitly tackles the issue of inter-slice inconsistencies by reconstructing 3D image volumes while accounting for multiple noise levels. However, DDPET-3D primarily focuses on denoising low-dose PET images rather than addressing the CT-to-PET synthesis problem, leaving a gap in the field of 3D modality translation. Despite significant progress, existing methods still face challenges when extended to 3D medical imaging. Directly stacking 2D slices into a 3D volume often results in discontinuities along the z-axis, reducing the overall spatial consistency and diagnostic value of the generated images. Both RegGAN and CPDM, while effective in the 2D space, do not address this limitation. On the other hand, DDPET-3D improves inter-slice consistency but is not designed for CT-to-PET translation.

In this work, we address these limitations by proposing RegC2P, a GAN-based framework that integrates a registration module to adaptively align PET slices with CT inputs. Unlike prior 2D methods, RegC2P generates spatially consistent 3D PET volumes by performing slice-by-slice translation while refining the results using a 3D U-Net-based network. This enables our method to produce high-quality, diagnostically relevant 3D PET images, overcoming the challenges posed by inter-slice inconsistencies and misaligned data.

3. PROBLEM AND METHODOLOGY

3.1. Problem definition

In PET/CT imaging systems, synthesizing PET images from CT images is crucial for effective diagnosis and treatment planning, as PET images highlight regions with high Standardized Uptake Values (SUVs), which indicate metabolic activity and are essential for identifying pathologies such as cancer. The CT2PET translation task addresses this issue and falls under the category of 3D I2I translation. However, modeling and computing the transformation between a 3D CT volume and a 3D PET volume require significant processing resources and complexity. Therefore, we propose reducing the 3D CT2PET problem to a slice-by-slice 2D CT2PET task, where the generated 2D slices are stacked to form a 3D volume. Traditional 2D I2I methods often struggle with maintaining structural fidelity and accurately capturing high-SUV regions, which can lead to less reliable diagnostic outcomes. Additionally, misregistration between CT and PET images further complicates this process, potentially resulting in misleading outputs.

To address these challenges, we aim to develop a mapping function fgen that transforms a CT slice SCT into a synthetic PET slice SPET, satisfying the following condition:

$$f_{gen}(S_{CT}) \approx S_{PET}, \quad (1)$$

with a focus on preserving anatomical structure and accurately representing high-SUV regions. This objective can be formulated using the following loss function:

$$L = \| f_{gen}(S_{CT}) - S_{PET} \|_p, \quad (2)$$

where $\| f_{gen}(S_{CT}) - S_{PET} \|_p$ enforces structural fidelity and accurate representation of SUV regions by minimizing the L_p-norm of the difference between the synthesized PET slice and the target PET slice. The parameter p is commonly set to either 1 (L1-norm) or 2 (L2-norm).

Inspired by [10], to further address the challenge of misregistration between CT and PET slices, we propose incorporating a regularization technique into the training process of a GAN-based model. This regularization aims to mitigate the impact of spatial misalignments. Our overview architecture is presented in figure 1.

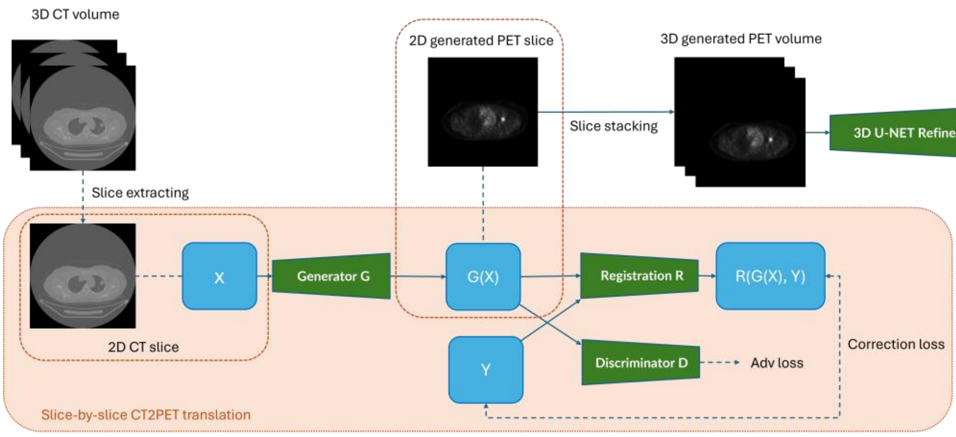

Figure 1. *RegC2P overview architecture.*

3.2. Registration regularization

In this study, we address the challenge of misaligned PET slices in CT2PET translation by treating them as *noisy labels*. Consequently, the CT2PET translation task is formulated as a supervised learning problem with noisy labels. Specifically, we are given a training dataset $\{(x_n, \tilde{y}_n)\}_{n=1}^{N}$, where x_n denotes a CT slice, and $\tilde{y}_n$ represents the corresponding misaligned PET slice. We assume y_n is the well-aligned PET slice (correct label) for x_n, which remains unknown in real-world scenarios.

Our objective is to train a generator G using the noisy dataset $\{(x_n, \tilde{y}_n)\}_{n=1}^{N}$ such that its performance approximates that of a generator trained on the clean dataset $\{(x_n, y_n)\}_{n=1}^{N}$. To achieve this, we adopt a *loss-correction* approach inspired by previous works [10]. The loss function is formulated as a log-likelihood to account for label noise, defined as:

$$L(G, R) = -\sum_{n=1}^{N} \log \left(p(\tilde{y}_n | y_n; R) p(\tilde{y}_n | x_n; G)\right) = -\sum_{n=1}^{N} \log \left(p(\tilde{y}_n | x_n; R, G)\right), \quad (3)$$

where R is a *registration network* that explicitly models the label noise as part of the network architecture. Here, the correct label y_n is treated as a *latent random variable*, enabling the model to learn the well-aligned PET distribution while accounting for noise in the training labels. The noise distribution in this problem can be modeled as a displacement error expressed as $\tilde{y} = y \circ D$, where D is a random *deformation field* that introduces spatial misalignment. To address this issue, we employ a registration network R after the generator G, which acts as a label noise model to correct the results. The correction loss is defined as:

$$\min_{G,R} L_{Corr}(G,R) = E_{x,\tilde{y}}[||y - G(x) \circ R(G(x),\tilde{y})||_1], \quad (4)$$

where $R(G(x),\tilde{y})$ is the deformation field, and $\circ$ denotes the resampling operation. The registration network R is implemented using a U-Net architecture [13]. To ensure smoothness in the deformation field R, we incorporate a smoothness loss, defined as:

$$\min_{R} L_{Smooth}(R) = E_{x,\tilde{y}}[||\nabla R(G(x),\tilde{y})||^2], \quad (5)$$

where ∇R represents the gradient of the deformation field, penalizing abrupt changes and enforcing smooth deformations. Additionally, we include the adversarial loss L_{Adv} to encourage realistic PET-like outputs in the GAN-based architecture. The overall objective combines the correction loss, smoothness loss, and adversarial loss, expressed as:

$$\min_{G,R} \max_{D} L_{Smooth}(G,R,D) = L_{Corr} + L_{Adv} + L_{Smooth}, \quad (6)$$

This formulation ensures that the generator G produces well-aligned PET slices while simultaneously addressing the effects of spatial misalignment and label noise through the registration network R and adversarial training.

3.3. 3D refinement network

After obtaining the generated 2D PET slices from the proposed GAN-based architecture in the previous section, we stack them to form a 3D PET volume. However, this stacking process introduces issues related to smoothness and consistency between consecutive slices, leading to artifacts such as discontinuities along the patient's body. These artifacts reduce the realism and diagnostic value of the resulting 3D PET images. To address this, we propose a 3D refinement network f_{ref} based on a 3D U-Net architecture.

The 3D refinement network f_{ref} takes as input a sub-volume $V_{PET_{pred}}$, consisting of *s* consecutive predicted PET slices extracted from the set of stacked 2D PET slices. The network outputs a corresponding refined sub-volume $V_{PET_{refined}}$, as described in the following equation:

$$V_{PET_{refined}} = f_{ref}(V_{PET_{pred}}), \quad (7)$$

where $V_{PET_{refined}}$ is the refined version of *s* consecutive slices.

To balance computational complexity and output quality, we fine-tune the hyperparameter *s* and evaluate the performance of the refined PET slices. Finally, to refine the entire 3D PET volume obtained from the previous phase, we propose an iterative refinement algorithm that processes sub-volumes of *s* consecutive slices, where each sub-volume is allowed to **overlap** with adjacent ones using a specific offset. The overlapping slices are aggregated by taking their average to ensure smooth transitions across slices, thereby improving the consistency and quality of the final 3D PET volume.

4. RESULTS AND DISCUSSION

4.1. Dataset

We introduce the 3D PET/CT dataset, which comprise 1,000 paired 3D CT-PET images. The dataset is meticulously assembled to support the advancement and assessment of cutting-edge medical image translation methods, providing a significant resource for researchers in the field. We then split the dataset into training, validation, and test sets using an 80:10:10 ratio, i.e., the training set consists of 80% of the studies. Preprocessing steps include standardizing all image pairs to a resolution of 256 × 256 × 1 and normalizing pixel values to the range of [−1, 1].

4.2. Results

In this section, we present both quantitative and qualitative evaluations of our proposed model against state-of-the-art methods for CT2PET translation. The performance is assessed using three key metrics: **Mean Absolute Error (MAE)**, **Structural Similarity Index (SSIM)**, and **Peak Signal-to-Noise Ratio (PSNR)**. Table 1 summarizes the quantitative results.

Table 1. *Performance comparison of CT2PET models.*

Model	**MAE ↓**	**SSIM ↑**	**PSNR ↑**
RegGAN	336.66	0.9241	28.40
DDPET-3D	351.00	0.9227	28.27
CPDM	402.85	0.9046	27.16
RegC2P (non-overlap)	**306.09**	**0.9317**	**29.32**
RegC2P (overlap)	**302.43**	**0.9330**	**29.42**

Lower (↓) MAE and higher (↑) SSIM/PSNR indicate better performance. The best and second-based results are highlighted by **red** and **blue**, respectively.

4.2.1. Quantitative results

Overall Performance: The proposed model, specifically the **RegC2P (over-lap)** variant, exhibits superior performance across all evaluated metrics. It achieves the **lowest MAE** (302.43), alongside the **highest SSIM** (0.9330) and **highest PSNR** (29.42), underscoring its efficacy in generating high-quality PET images from CT inputs with remarkable accuracy and fidelity.

Impact of Overlapping Slices: The results highlight the significant benefits of incorporating overlapping slices during iterative refinement, leading to marked improvements in model performance across all evaluation metrics. Specifically, the inclusion of overlapping slices reduces the MAE from 306.09 in the non-overlapping variant (**RegC2P non-overlap**) to 302.43 in the overlapping version (**RegC2P overlap**), demonstrating enhanced accuracy in the generated PET images. Additionally, the SSIM shows a noticeable increase, improving from 0.9317 to 0.9330, reflecting better preservation of structural and textural fidelity. Furthermore, the PSNR exhibits a slight yet meaningful rise, advancing from 29.32 to 29.42, indicating reduced noise and sharper image quality. These results underscore the effectiveness of overlapping slice refinement in improving both quantitative metrics and the perceptual quality of generated images.

4.2.2. Qualitative results

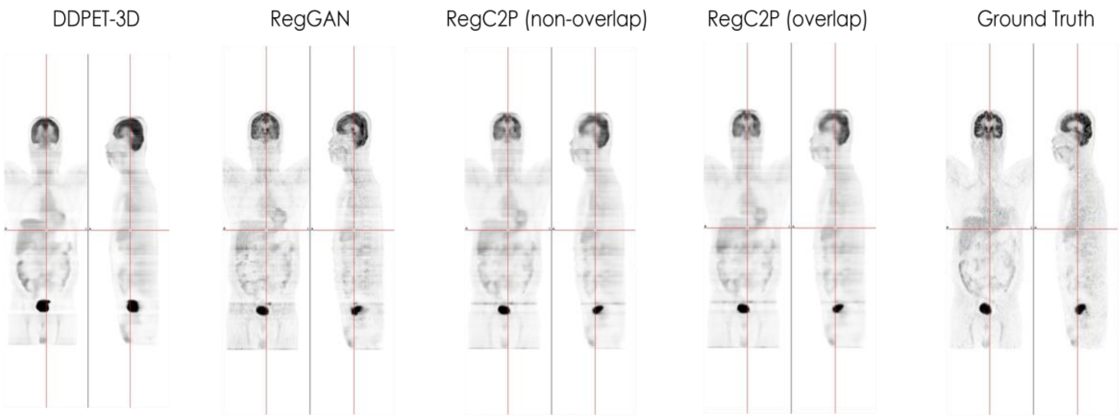

Figure 2. *Qualitative comparison of generated PET volumes. From left to right: DDPET-3D, RegGAN, RegC2P (non-overlap), RegC2P (overlap) and ground truth PET.*

To further evaluate the performance of our method, we provide visual comparisons of the generated PET volumes with the ground truth and results from baseline methods. Figure 1 shows qualitative examples of PET slices generated by different models.

The RegC2P (overlap) model demonstrates a superior ability to generate PET slices that closely align with the ground truth, characterized by smooth transitions and accurately preserved high-SUV regions. In comparison to RegGAN and DDPET-3D, our model exhibits enhanced structural coherence, with a significant reduction in artifacts and improved sharpness of fine details. While the non-overlapping variant, RegC2P (non-overlap), produces images of commendable quality, it displays minor boundary inconsistencies. These artifacts are effectively resolved through the overlapping refinement process, highlighting the robustness of the proposed approach.

5. CONCLUSIONS

The experimental results, both quantitative and qualitative, clearly demonstrate the effectiveness of our proposed method. Our model not only surpasses state-of-the-art baselines in terms of accuracy and image quality but also produces visually consistent and diagnostically relevant PET images. The overlapping slice refinement approach plays a crucial role in improving smoothness and structural consistency in the synthesized 3D PET volumes.

To further elevate the performance of RegC2P, future research will explore the development of hybrid models that integrate GAN with diffusion techniques. In particular, we aim to adopt the I2I diffusion model as a baseline for generating 2D slices instead of GAN-based model. This approach seeks to capitalize on the exceptional sharpness and realism provided by diffusion models, thereby addressing existing limitations and advancing model capabilities. Furthermore, we substitute the 3D U-NET with a GAN-based refinement model to further elevate the fidelity and quality of the synthesized 3D PET images. By leveraging the strengths of both approaches, we aim to achieve a superior balance between the visual realism of synthesized images and their pixel-wise accuracy, which is crucial for precise medical interpretation. These efforts are aligned with the overarching goal of refining RegC2P into a highly reliable and clinically applicable tool for PET image synthesis. Through these initiatives, RegC2P aspires to contribute meaningfully to the field of medical imaging by providing innovative diagnostic tools that support better clinical decision-making, improve patient outcomes, and drive advancements in healthcare technology.

REFERENCES

[1]. L. Wei and I. El Naqa, *"Artificial intelligence for response evaluation with PET/CT,"* in Seminars in Nuclear Medicine, vol. 51, pp. 157–169, Elsevier, (2021).

[2]. S. R. Cherry, T. Jones, J. S. Karp, J. Qi, W. W. Moses, and R. D. Badawi, *"Total-body PET: maximizing sensitivity to create new opportunities for clinical research and patient care,"* Journal of Nuclear Medicine, vol. 59, no. 1, pp. 3–12, (2018).

[3]. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, *"Generative adversarial nets,"* Advances in Neural Information Processing Systems, vol. 27, (2014).

[4]. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, *"Image-to-image translation with conditional adversarial networks,"* in Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1125– 1134, (2017).

[5]. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, *"Unpaired image-to-image translation using cycle-consistent adversarial networks,"* in Proceedings of the 2017 IEEE International Conference on Computer Vision, pp. 2223–2232, (2017).

[6]. H. Tang, D. Xu, N. Sebe, Y. Wang, J. J. Corso, and Y. Yan, *"Multichannel attention selection GAN with cascaded semantic guidance for cross-view image translation,"* in Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2417– 2426, (2019).

[7]. K. Armanious, C. Jiang, M. Fischer, T. Kustner, T. Hepp, K. Nikolaou, S. Gatidis, and B. Yang, *"Medgan: Medical image translation using GANs,"* Computerized Medical Imaging and Graphics, vol. 79, p. 101684, (2020).

[8]. Z. Liang, H. Wei, G. Liu, M. Cheng, J. Gao, S. Li, and X. Tian, *"Leveraging GAN-based CBCT-to-CT translation models for enhanced image quality and accurate photon and proton dose calculation in adaptive radiotherapy,"* Journal of Radiation Research and Applied Sciences, vol. 17, no. 1, p. 100809, (2024).

[9]. Y. Zhu, Y. Tang, Y. Tang, D. C. Elton, S. Lee, P. J. Pickhardt, and R. M. Summers, *"Cross-domain medical image translation by shared latent gaussian mixture model,"* in Proceedings of the 23rd International Conference Medical Image Computing and Computer Assisted Intervention, pp. 379–389, Springer, (2020).

[10]. L. Kong, C. Lian, D. Huang, Y. Hu, Q. Zhou, et al., *"Breaking the dilemma of medical image-to-image translation,"* Advances in Neural Information Processing Systems, vol. 34, pp. 1964–1978, (2021).

[11]. D. T. Nguyen, T. T. Nguyen, H. T. Nguyen, T. T. Nguyen, H.-H. Pham, T. H. Nguyen, T. N. Truong, and P. L. Nguyen, *"CT to PET Translation: A Large-scale Dataset and Domain-Knowledge-Guided Diffusion Approach,"* Accepted for publication at the 2025 IEEE/CVF Winter Conference on Applications of Computer Vision, (2024).

[12]. Xie, Huidong, et al. *"DDPET-3D: Dose-aware Diffusion Model for 3D Ultra Low-dose PET Imaging." arXiv preprint arXiv:2311.04248* (2023).

[13]. O. Ronneberger, P. Fischer, and T. Brox, *"U-net: Convolutional networks for biomedical image segmentation,"* in Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 234–241, (2015).

TÓM TẮT

RegC2P: Một mô hình sinh ảnh 3D-PET từ ảnh 3D-CT dựa trên GAN

Sự kết hợp giữa phương pháp Chụp cắt lớp phát xạ Positron (PET) và Chụp cắt lớp vi tính (CT) là bước tiến quan trọng trong lĩnh vực chẩn đoán ung thư bằng cách kết hợp thông tin chuyển hóa và giải phẫu, giúp cải thiện độ chính xác trong chẩn đoán, phân giai đoạn và theo dõi điều trị bệnh. Tuy nhiên, việc ứng dụng rộng rãi hệ thống PET/CT vẫn còn hạn chế do sự khan hiếm của máy quét PET và phương pháp chụp PET phụ thuộc vào phóng xạ. Để khắc phục vấn đề này, bài toán dịch ảnh 3D nổi lên như một giải pháp tiềm năng nhằm tạo ra ảnh PET tổng hợp từ ảnh CT. Tuy nhiên, các phương pháp sinh ảnh hiện có dựa trên Mạng sinh đối nghịch (GANs) gặp phải những thách thức như sự bất ổn định và ngẫu nhiên trong kết quả sinh, khiến quá trình chuyển đổi CT sang PET trong không gian 3D chưa đạt được độ chính xác mong muốn. Chúng tôi đề xuất RegC2P, một phương pháp mới tích hợp mô-đun đăng ký ảnh vào kiến trúc GAN để tạo ra ảnh PET 3D được căn chỉnh chính xác từ ảnh CT. RegC2P tiếp cận bài toán này theo hướng dịch ảnh 2D theo từng lát cắt, trong đó, mỗi lát cắt CT 2D được chuyển đổi thành lát cắt PET 2D tương ứng, sau đó xếp chồng để tạo thành ảnh PET 3D. Các lát cắt PET 2D bị căn chỉnh lệch được xem như nhãn nhiễu, và mạng sinh ảnh của RegC2P được huấn luyện kết hợp với mạng đăng ký để tự động hiệu chỉnh sai lệch, tối ưu hóa đồng thời cả hai nhiệm vụ dịch ảnh và căn chỉnh ảnh. Mặt khác, để đảm bảo độ mượt và tính nhất quán giữa các lát PET 2D được sinh ra trong toàn bộ khối ảnh 3D, chúng tôi đề xuất một mạng tinh chỉnh 3D U-Net. Các thí nghiệm trên tập dữ liệu lớn cho thấy RegC2P vượt trội hơn các phương pháp tiên tiến hiện nay, đạt giảm 10.16% MAE, tăng 0.96% SSIM và tăng 3.6% PSNR, từ đó thiết lập một tiêu chuẩn mới về chất lượng ảnh PET 3D tổng hợp từ mô hình Học sâu.

Từ khóa: Dịch ảnh 3D; Chụp xạ hình 3D PET/CT; Chuyển ảnh CT sang PET; GAN; U-Net.