

# Non-complete-ring positron emission tomography (PET) detection method

Yeqi Fang<sup>1,\*</sup> and Rong Zhou<sup>1,†</sup>

<sup>1</sup>College of Physics, Sichuan University, Chengdu, 610065, China

Positron Emission Tomography (PET) is a vital molecular imaging tool widely used in medical diagnosis and treatment evaluation. Traditional PET systems typically rely on complete detector rings to achieve full angular coverage for uniform and statistically robust sampling of coincidence events. However, incomplete-ring PET scanners have emerged in various scenarios due to hardware failures, cost constraints, or specific clinical needs. In such cases, conventional reconstruction algorithms often suffer from performance degradation due to reduced data completeness and geometric inconsistencies. This thesis proposes a coarse-to-fine reconstruction framework for incomplete-ring PET scanners. The framework first employs an Attention U-Net model to recover complete sinograms from incomplete ones, then uses the OSEM algorithm for preliminary reconstruction, and finally applies a two-stage architecture comprising a Coarse Prediction Module (CPM) and an Iterative Refinement Module (IRM) for fine reconstruction. Our approach utilizes neighboring axial slices and spectral transform features as auxiliary guidance at the input level to ensure spatial and frequency domain consistency, and integrates a contrastive diffusion strategy at the output level to improve correspondence between low-quality PET inputs and refined PET outputs. Experimental results on public and in-house brain PET datasets demonstrate that the proposed method significantly outperforms existing approaches in metrics such as PSNR (35.6421 dB) and SSIM (0.9588), successfully preserving key anatomical structures and tracer distribution features, thus providing an effective solution for incomplete-ring PET imaging.

## I. INTRODUCTION

### A. Background and Motivation

Positron emission tomography (PET) is a powerful molecular imaging technique that provides quantitative visualization of metabolic processes within living tissue [1]. The fundamental principle is elegant but complex: a radiotracer introduces positrons that annihilate upon encountering electrons, and then produce pairs of 511keV photons traveling in opposing directions. The photon pairs are then detected by a pair of scintillation sensors, and each coincidence event will define a line-of-response (LOR). Traditional PET systems usually use complete 360° detector rings to maximize their sensitivity and angular coverage [1].

But what happens when a complete ring isn't feasible? This question has driven the development of incomplete-ring PET systems. Such systems have emerged from practical necessity rather than theoretical preference: they reduce costs and complexity, allow closer access to patients in specialized applications like breast scanning [2], create "open" configurations that alleviate claustrophobia [3, 4], and enable novel applications such as dual-panel brain imaging systems [5]. The trade-off is severe—complete angular coverage creates gaps in projection data, turning reconstruction into an underdetermined problem. Missing angular views inevitably introduce artifacts and resolution non-uniformity [2, 6]. Even time-of-flight capabilities, which partially fix these problems, can't fully compensate for limited view angles [2, 4].

Researchers have investigated a lot of ways to achieve better and better incomplete-ring PET reconstruction. Analytical ways like filtered back-projection falter immediately—their assumption of complete data leads to pronounced streak artifacts [6]. Iterative approaches like maximum-likelihood expectation-maximization (MLEM) are better by modeling the acquisition process [7], but they still have artifacts along missing view directions [5]. Penalized likelihood methods incorporating prior constraints show promise, as Zhang *et al.* demonstrated by using high-quality prior images to enhance contrast recovery in a dual-panel head-and-neck PET system [5]. But these methods are still not enough to solve the fundamental problem.

Deep learning has brought new ideas to this challenge, though with mixed results. Regression-based networks have shown promise in low-dose PET denoising and partial data reconstruction [8] by learning direct mappings between degraded and high-quality images. Liu *et al.*'s U-Net approach to transform artifact-degraded partial-ring PET images [9] demonstrated initial success, but conventional CNNs struggle with the global features needed to address large-scale angular losses.

Current available techniques have some drawbacks. For examples, while GAN approaches are good at generating visually compelling reconstructions, [10] they are unstable when training and often highly sensitive to hyperparameters, which limit their clinical use. Methods based on explicit likelihood modeling—like VAEs and flow-based approaches—offer solid theoretical foundations. But they usually lose details in reconstructed images and their low speeds also make them not suitable for clinical use. Model-based deep learning frameworks and generative models [11, 12] show promise but risk introducing hallucinated features—a dangerous proposi-

---

\* fangyeqi@stu.scu.edu.cn

† zhourong@scu.edu.cn

tion in medical imaging.

The core challenge is the fundamental **information loss**, which can not be solved, only mitigation. Analytical methods collapse under data deficiency, over-regularized iterative methods blur subtle features, and deep learning models risk introducing misleading artifacts. Current approaches fail in at least one critical dimension—they can't handle severe data loss, require unrealistic computational resources, or compromise the detailed features that clinicians depend on for diagnosis. This study solves these problems by introducing a new coarse-to-fine model for both sinogram and its reconstruction which can preserve details useful for accurate clinical diagnosis.

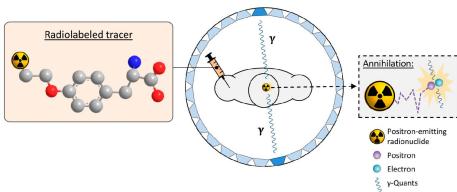


FIG. 1. Principles of Positron Emission Tomography (PET) imaging.

The remainder of this thesis is organized as follows: **Section II** (*Background on PET Imaging and Incomplete Ring Geometry*) provides an overview of PET and its basic physics like principles of coincidence detection, and the main problems caused by incomplete rings. We also summarize classical and modern reconstruction methods. **Section III** (*Fundamentals of Diffusion Probabilistic Models*) details the background of Denoising Diffusion Probabilistic Models (DDPM) and how they apply to various image reconstruction tasks, establishing the theoretical foundation for our proposed method. *Proposed Coarse-to-Fine Reconstruction Framework* describes the detailed workflow of our method, including complete and incomplete ring geometric modeling, creation of list-mode and sinogram data, the coarse-to-fine design, auxiliary guidance modules, contrastive diffusion learning objectives, and implementation details. **Section IV** (*Experiments and Results*) presents our experimental setups including the restoration of both sinogram and its reconstruction. We also include hyperparameter choices, metrics to assess the performance of our model, and comparisons with other excellent methods. **Section V** (*Summary and Prospective*) summarizes our findings by analysing metrices of model performance on test dataset, and discusses limitations and directions for future improvements. **Appendices** provide additional experiments, extended qualitative results, and tables of hyperparameters used in the study.

## II. BACKGROUND ON PET IMAGING AND INCOMPLETE RING GEOMETRY

When two detectors measure gamma photons within a narrow time window (typically 6-12 nanoseconds), we assume they are from the same annihilation event. This paired detection forms a line of response (LOR) between the two detectors. PET systems record these events in one of two primary formats. In **list-mode data**, each coincidence is logged individually with precise identification of the detector pair and timestamp. This approach saves the maximum amount of information but generates enormous datasets. **Sinogram data** organizes coincidence events into radial, angular, and axial bins, creating a structured histogram that trades off temporal resolution for simpler processing. Standard PET scanners use 360° detector rings to achieve uniform LOR sampling, but things would be quite different when this ideal geometry isn't possible.

### A. Incomplete Ring PET Scanners

Incomplete ring configurations emerge from a clash between ideal physics and practical reality. Budget constraints often drive this compromise—each detector module constitutes a substantial cost, and reducing their number can make PET technology accessible to more facilities. In specialized applications, partial ring designs actually provide benefits: breast imaging benefits from closer detector positioning, interventional procedures require open scanner designs, and claustrophobic patients experience less anxiety with more open configurations.

The consequences of missing detectors are profound. LORs that would normally pass through the missing sections simply vanish from the dataset, creating angular sampling gaps. We will see a mess if we directly reconstruct the image out of incomplete sinogram, as shown in Figure 3.

### B. System Model and Data Simulation

To deal with these problems systematically, we made a thorough simulation framework modeling both complete and incomplete PET geometries. Our approach begins with a standard ring configuration of  $R$  axial rings (we used  $R = 42$  in our experiments), each containing  $D = 182$  detectors arranged cylindrically. The scanner radius ( $\rho$ ) of 253.71 mm balances spatial resolution against sensitivity.

Early in our research, we attempted to use Transformer-based approaches for direct 3D image recovery, following Hatamizadeh et al.'s UNETR framework [13]. The segmentation results were excellent, but when applied to our incomplete sinogram data, these methods did not work well, with SSIM of less than 0.2 even after a long time of training. This indicates that mere geometric

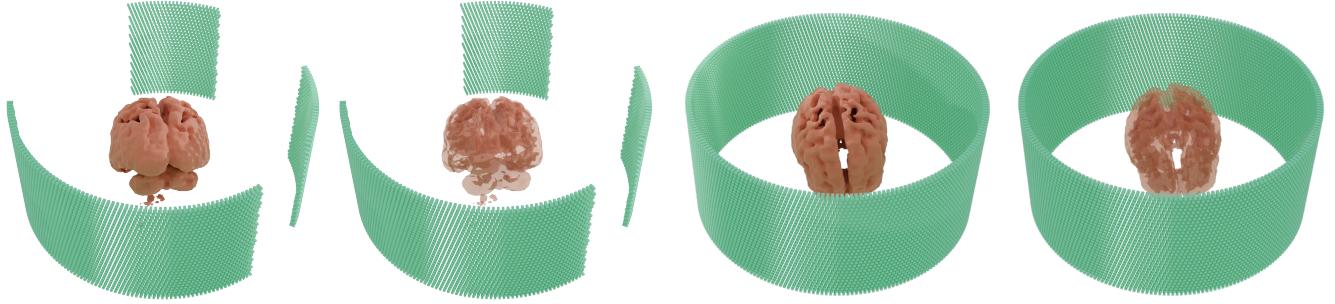


FIG. 2. Three-dimensional schematic of incomplete and complete ring PET scanner detector structure, with the right image showing a perspective view.

information is insufficient for high-quality reconstruction when some angular segments are missing.

From listmode to sinogram, and from sinogram to reconstruction, the information is constantly losing. So, rather than attempting direct image-domain recovery, we pivoted to a two-stage method—first converting listmode data to sinograms with missing sections, then applying specialized techniques to complete these sinograms before final reconstruction. This approach, while computationally more intensive, provided a more applicable way to solve the missing data problem by retaining as much information as possible from sinogram. For simulation purposes, we mapped a  $128 \times 128 \times 128$  voxel grid to physical space, yielding in-plane voxel dimensions of approximately 2.78 mm.

The detector system itself employs a hierarchical structure with parameters that were hard to optimize. Table I describes some important configuration parameters of our PET system.

TABLE I. PET Scanner Configuration Parameters

Parameter	Value
Radius	253.71 mm
Crystal transaxial spacing	4.02 mm
Crystal axial spacing	5.37 mm
Module axial spacing	37.56 mm
Crystal elements (transaxial)	13
Crystal elements (axial)	7
Transaxial sectors	28
Axial sectors	1
Modules (axial)	6
Crystals per ring	364
Number of rings	42

Our theoretical model parameters is an idealized PET system, setting aside real-world manufacturing variations that would typically require complex normalization procedures. While our framework includes parameters for TOF detection, we chose not to implement this capability in the current work, because the devices to which we intend to apply this study do not have that high temporal resolution.

Similarly, our simulation framework does not account for attenuation correction problems needing MRI or CT

data in physical implementations—an important consideration for future translation of our approach. Our simulation generates idealized datasets by ray-tracing approximately 20 million emission events from  $128^3$ -voxel phantom images, creating perfectly clean list-mode data without the noise and artifacts present in real-world situation. We used a dataset of 206 brain PET volumes, each with dimensions of  $128^3$ -voxel from a public dataset as phantom images for our simulation [44]. Each image simulates 2 billion events, the detectors detect 560 million events on average and record them in the list-mode data. We then removed detectors with angles between  $[30^\circ, 90^\circ]$  and  $[210^\circ, 270^\circ]$ . This meant that any LOR with either detector within these two angular ranges cannot be recorded. Incomplete sinograms generated in this configuration only have about 306 million events, accounting for 54.6% of the originally detected events.

The complete ring configuration (Figure 2) provides our reference reconstruction  $\mathbf{Y}_A$ , representing the best-case scenario with full angular sampling. For incomplete ring experiments (Figure 2), we selectively remove detectors—either entire rings or angular segments—and generate a degraded reconstruction  $\mathbf{Y}_B$  from the resulting partial data. The striking quality difference between  $\mathbf{Y}_A$  and  $\mathbf{Y}_B$  shows the reconstruction challenge we aim to solve, as shown in Figure 3.

### C. Conventional PET Reconstruction Methods

Before introducing new solutions, Several popular reconstruction approaches that are not based on machine learning have been assessed. Currently, main powerful methods are **maximum likelihood methods**, particularly MLEM and its faster variant OSEM [14], which include the Poisson statistics of photon counting. The most important parameter of these two approaches is **system model**, which relates image voxels to measured projections:

$$y_i \approx \sum_j p_{ij} \lambda_j, \quad (1)$$

Here,  $\lambda_j$  is the activity in voxel  $j$ ,  $y_i$  is the count measured in projection  $i$ , and  $p_{ij}$  is the probability that an

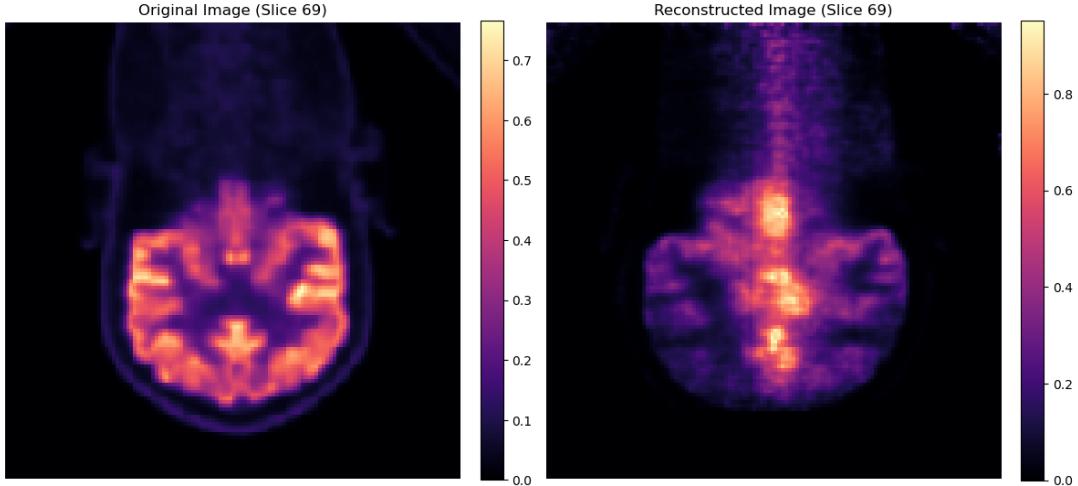


FIG. 3. Comparison of reconstruction results from incomplete rings with the original image, showing that the two are vastly different, demonstrating that direct reconstruction from incomplete rings is not feasible.

emission from voxel  $j$  is detected along projection  $i$ . This **system matrix**  $p_{ij}$  includes all the physics—from detector geometry to attenuation effects—that connects the unknown tracer distribution to our measurements.

MLEM iteratively updates the estimate of  $\lambda_j$  according to:

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} \times \frac{\sum_{i=1}^N \frac{p_{ij}}{\sum_\ell p_{i\ell} \lambda_\ell^{(k)}} y_i}{\sum_{i=1}^N p_{ij}}. \quad (2)$$

However, MLEM’s convergence is too slow for large datasets. OSEM makes this process faster by dividing projections into subsets and then update the image estimate using one subset at a time:

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} \times \frac{\sum_{i \in S_k} \frac{p_{ij}}{\sum_\ell p_{i\ell} \lambda_\ell^{(k)}} y_i}{\sum_{i \in S_k} p_{ij}} \quad (3)$$

Our work mainly depend on the PyTomography library [15], which make use of PyTorch and GPUs to make intense computations a lot faster. This technique is important when reconstructing large listmode data with around one billion events. While OSEM is powerful with full angular sampling, its performance with incomplete rings was disappointing [cf. Figure 3]. Because the missing angles also mean information loss that makes any iterative methods impossible to recover, despite their statistical sophistication. Thus, OSEM can only be useful if the sinogram is complete or recover to a complete one. But it can not solve the core problem of missing data.

### III. IMPLEMENTATION METHODS

As shown in Figure 4, this study proposes an innovative incomplete ring PET reconstruction framework that effectively addresses the data incompleteness problem caused by missing detectors through a multi-stage strategy. In the first stage, the system first processes incomplete sinograms (top left) resulting from missing detectors, completing the missing data through a trained optimized U-Net deep learning model to generate complete sinograms (top right). This sinogram restoration process fully utilizes the five-channel input strategy described in Chapter 3, effectively integrating spatial and temporal context information.

#### A. Sinogram Reconstruction Based on Attention U-Net

Sinogram repair is very important in incomplete ring PET imaging, as its quality directly affects the precision of reconstructed images. The shape of sinogram tensors is  $(D, 2D + 1, R^2)$ , or  $(182, 365, 1764)$  in our work, where  $R$  is the number of axial rings and  $D$  is the number of detectors each ring. And it is obvious that this tensor is too large to be directly fed into any UNet-like models. So we cut this large  $(D, 2D + 1, R^2)$  tensor into  $R \times (D, 2D + 1, R)$  tensors. As shown in Figure 5, the central channel of each input tensor corresponds to the current sinogram slice, while the direct spatial neighbors (slices  $j - 1$  and  $j + 1$ ) and temporal neighbors from previous and subsequent sinogram periods (slices  $j - R$  and  $j + R$ ) constitute the other four channels. For boundary handling, when adjacent indices exceed the dataset range, the central slice itself is used for channel filling, ensuring input dimension consistency. These neighboring slices are important for improving the model’s understanding of local structural continuity and temporal

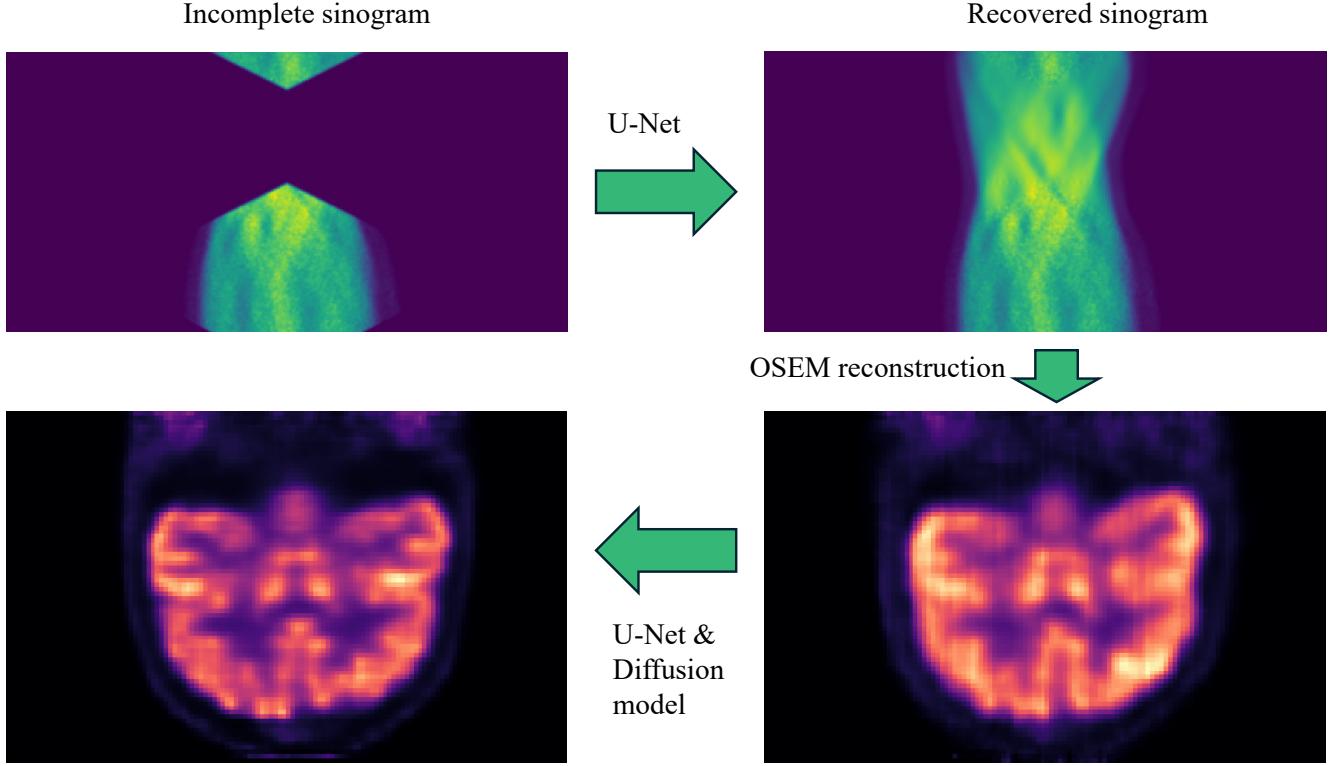


FIG. 4. Overall flow diagram of the incomplete ring PET reconstruction method. First, incomplete sinograms due to missing detectors (top left) are repaired through a U-Net deep learning model to generate complete sinogram data (top right). Subsequently, the complete sinogram is converted into a PET image (bottom right) using the OSEM iterative reconstruction algorithm. For comparison, the bottom left image shows a PET image reconstructed directly from incomplete data using a coarse-to-fine framework (dual-model architecture with U-Net and diffusion model).

consistency. So finally we have  $1764 \times (182, 365, 5)$  tensors for each sinogram. To make sure each image was evaluated once in the test set, we implemented 6-fold cross-validation across the 206 volumes (4 groups with 34 images and 2 groups with 35 images). The predictions obtained from our model were reconstructed using the OSEM algorithm to achieve preliminary geometric recovery, and these reconstructions were subsequently used as inputs for our the coarse-to-fine framework.

Although traditional U-Net performs excellently in medical image segmentation and reconstruction tasks [16], it lacks the ability to selectively focus on key features when dealing with complex incomplete ring PET sinogram restoration problems. The Attention U-Net [17] adopted in this study enhances the model's perception of important feature regions through spatial attention mechanisms while suppressing the influence of irrelevant features, which is more useful for restoring sinograms from incomplete data. We also tested some transformer-based U-Net like UNETR [13], which combines ViT encoder and convolutional decoder, and TransUNet [18], where both encoder and decoder are completely based on transformers. But they are slightly outperformed by Attention U-Net.

The main innovation of Attention U-Net is the introduction of attention gating (AG) modules in the skip

connection path of the original U-Net. These AG modules can adaptively highlight significant structures in the feed-forward feature maps while suppressing less relevant regions. For sinogram reconstruction tasks, this mechanism is especially useful as it can selectively focus on structural features around missing areas, and then more accurately infer missing angular data. The mathematical expression of attention gating can be described as [17]:

$$\alpha_i^l = \sigma_2(\psi^T(\sigma_1(W_x^T x_i^l + W_g^T g_i + b_g)) + b_\psi), \quad (4)$$

where  $x_i^l$  is the low-level feature from the encoder,  $g_i$  is the gating signal from the decoder (high-level feature),  $\sigma_1$  and  $\sigma_2$  are ReLU and Sigmoid activation functions respectively, and  $W_x$ ,  $W_g$ ,  $b_g$ , and  $b_\psi$  are learnable parameters.  $\alpha_i^l \in [0, 1]$  is the calculated attention coefficient used to control the importance of features. After processing through the attention gate, the features can be represented as:

$$\hat{x}_i^l = x_i^l \cdot \alpha_i^l \quad (5)$$

The Attention gates learn to focus on relevant regions of the encoder feature maps by assigning weights based on the context, particularly in cases of severe angular loss. AGs can create a stronger understanding of sinogram continuity and consistency.

After testing on our dataset, the results showed that Attention U-Net performs better with an average increase of 1.24dB in PSNR and 0.052 in SSIM compared to original U-Net in sinogram reconstruction tasks. This means that attention mechanism is more effective in processing incomplete ring PET data.

TABLE II. Five-channel U-Net network structure parameters

Description	Type	Kernel	Output
Input	-	-	5
Encoder Block 1	Conv-BN-ReLU	$3 \times 3$	64
Pooling	Max Pooling	$2 \times 2$	64
Encoder Block 2	Conv-BN-ReLU	$3 \times 3$	128
Pooling	Max Pooling	$2 \times 2$	128
Encoder Block 3	Conv-BN-ReLU	$3 \times 3$	256
Pooling	Max Pooling	$2 \times 2$	256
Encoder Block 4	Conv-BN-ReLU	$3 \times 3$	512
Pooling	Max Pooling	$2 \times 2$	512
Bottleneck	Conv-BN-ReLU	$3 \times 3$	1024
Decoder Block 1	Upsampling + Conv-BN-ReLU	$3 \times 3$	512
Decoder Block 2	Upsampling + Conv-BN-ReLU	$3 \times 3$	256
Decoder Block 3	Upsampling + Conv-BN-ReLU	$3 \times 3$	128
Decoder Block 4	Upsampling + Conv-BN-ReLU	$3 \times 3$	64
Output Layer	Conv	$1 \times 1$	5

The model structure is shown in Tabel II. Model training uses the Adam optimizer with an initial learning rate of  $10^{-4}$  and adds  $10^{-5}$  weight decay to prevent overfitting. Training efficiency is enhanced through mixed precision computation and gradient scaling techniques, combined with a ReduceLROnPlateau dynamic learning rate scheduler (with a decay factor of 0.3 when validation loss shows no improvement for 3 consecutive epochs), achieving stable convergence.

## B. Coarse-to-Fine Reconstruction Framework

After repairing incomplete sinogram to a complete one, the image quality is still not satisfactory with its SSIM lower than 0.9. To properly deal with this problem, we adopt a two-stage approach that can restore the image on the geometry space, which can further improve image quality.

### 1. Coarse Prediction Module (CPM) and Iterative Refinement Module (IRM)

The first stage employs a traditional U-Net as the coarse prediction module that generates an initial reconstruction estimate. Let  $\mathbf{c}$  represent our conditional inputs, comprising the low-quality PET image  $\mathbf{Y}_B$  from incomplete data and auxiliary guidance information  $\mathbf{X}_{\text{aux}}$ .

The CPM, implemented as a deterministic U-net, produces a coarse reconstruction  $\mathbf{x}_{\text{cp}}$  through a single forward pass:

$$\mathbf{x}_{\text{cp}} = P_{\theta}(\mathbf{c}) \quad (6)$$

This module can capture general image structures and give a good starting point for diffusion model later on [25]. This U-Net is also optimized using an L1 loss:

$$\mathcal{L}_{\text{CPM}} = \|\mathbf{x}_{\text{cp}} - \mathbf{Y}_A\|_1 \quad (7)$$

where  $\mathbf{Y}_A$  represents the target high-quality reconstruction.

The second stage uses a Diffusion probabilistic model (DPM) as iterative refinement module that focuses specifically on recovering residual details.

Diffusion models provide several advantages for PET reconstruction under the noisy and incomplete data conditions [21]. They can avoid the mode collapse problem common in GANs, making sure robust coverage of possible solutions and also reconstruct details that are useful for clinical diagnosis but usually no captured by U-Net [22]. They can also integrate learned priors for PET images, which is important when large segments of sinogram data are missing [23]. However, directly applying DPMs for reconstruction introduces issues pertaining to computation speed and partial data handling [9]. To fix these problems, we adopt a coarse-to-fine method that introduces a deterministic, high-capacity coarse prediction module combined with a smaller diffusion model focused specifically on residual reconstruction [24].

Diffusion models are a class of probabilistic generative models proposed by *Sohl-Dickstein et al.*[19], and later widely applied in image generation tasks by *Ho et al.*[20] and also medical images. Their framework can be understood through two main processes:

The forward process gradually adds noise to the data  $\mathbf{x}_0 \in \mathbb{R}^d$  through a Markov chain. At each step, Gaussian noise is added according to a predefined schedule:

$$q(\mathbf{x}_t | \mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1 - \beta_t} \mathbf{x}_{t-1}, \beta_t \mathbf{I}) \quad (8)$$

where  $\beta_t \in (0, 1)$  represents the noise intensity at step  $t$ . The process eventually transforms the data distribution into an isotropic Gaussian distribution.

The reverse process is the core of generation, where a neural network is trained to denoise gradually, starting from Gaussian noise:

$$p_{\theta}(\mathbf{x}_{t-1} | \mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_{\theta}(\mathbf{x}_t, t), \sigma_t^2 \mathbf{I}) \quad (9)$$

The model learns to predict the noise component added in the forward process, allowing for step-by-step recovery of the clean image.

The training objective is usually simplified to predicting the noise added in the forward process:

$$L(\theta) = \mathbb{E}_{\mathbf{x}_0, \epsilon, t} [\|\epsilon - \epsilon_{\theta}(\mathbf{x}_t, t)\|^2] \quad (10)$$

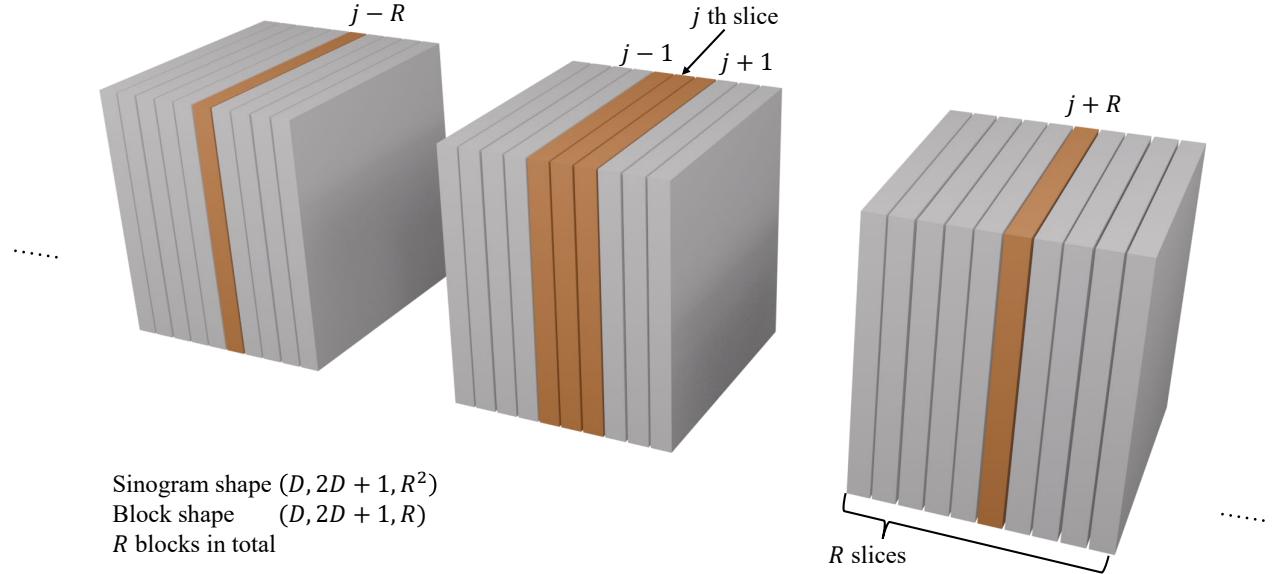


FIG. 5. Visualization of the five-channel tensor preparation of sinogram data for training. Each cube represents a sinogram slice block with dimensions  $(D, 2D + 1, R^2)$ . The orange highlighted parts are the selected slices  $(j - R, j - 1, j, j + 1, j + R)$ , which form the five-channel input for the restoration model, showing the spatial and temporal relationships captured in each input tensor.

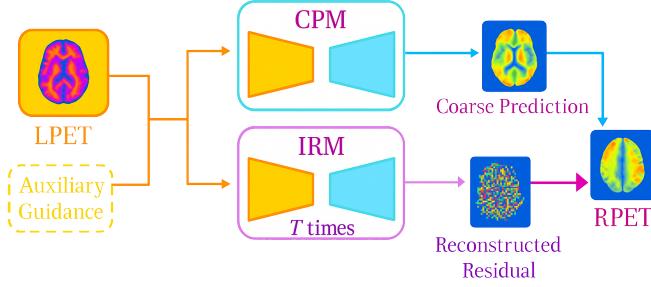


FIG. 6. Overview of the coarse-to-fine incomplete ring PET reconstruction framework. The method contains two key modules: the Coarse Prediction Module (CPM) and the Iterative Refinement Module (IRM). The system first receives low-quality PET images (LPET) and auxiliary guidance information as input. The CPM generates preliminary reconstruction results (coarse prediction) through a single forward pass. Subsequently, the IRM, based on diffusion probabilistic models, reconstructs the residual signal through  $T$  iterative steps. Finally, the system combines the coarse prediction with the reconstructed residual to generate a high-quality refined PET image (RPET). This dual-stage design effectively balances computational efficiency with reconstruction quality, overcoming the limitations of traditional diffusion models in terms of computational cost.

During sampling, we start from Gaussian noise  $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ , and then the model iteratively refines the sample through several denoising steps.

Rather than modeling the entire image, the IRM models only the residual  $\mathbf{r}_0 = \mathbf{Y}_A - \mathbf{x}_{cp}$ . This residual-focused methodology provides substantial benefits: Firstly, the IRM can utilize a smaller network capacity as it only

needs to capture fine details. Secondly, the distribution gap between coarse prediction and ground truth is substantially smaller than between random noise and ground truth. Lastly, this narrower gap enables faster convergence and fewer diffusion steps during inference.

The final reconstruction is obtained by combining the coarse prediction with the estimated residual:

$$\widehat{\mathbf{Y}}_A = \mathbf{x}_{cp} + \widehat{\mathbf{r}}_0 \quad (11)$$

## 2. Auxiliary Guidance Strategies

To further improve reconstruction performance in regions with some discrepancy between the prediction reconstructed from restored sinogram and the ground truth, we add two complementary guidance images:

**Neighboring Axial Slices (NAS).** Since PET volumes are 3-dimensional, we use adjacent slices to provide more spatial context [26], which is similar to what we have done in section III A:

$$\mathbf{X}_{NAS} = \{\mathbf{Y}_B^{(z-2)}, \mathbf{Y}_B^{(z-1)}, \mathbf{Y}_B^{(z+1)}, \mathbf{Y}_B^{(z+2)}\} \quad (12)$$

This multi-slice approach enables the model to maintain structural consistency across the axial dimension, reducing isolated artifacts that might appear in single-slice approaches.

**Spectral Guidance.** We also incorporate frequency domain information by applying the discrete Fourier transform to input slices:

$$\mathbf{X}_{spec} = \mathcal{F}(\mathbf{Y}_B) \quad (13)$$

Spectral guidance provides global frequency priors that help suppress the high-frequency streak artifacts commonly associated with missing information in PET reconstruction [27].

This coarse-to-fine framework effectively balances reconstruction quality with computational efficiency, providing a practical solution for incomplete ring PET imaging in clinical settings.

### C. Evaluation Metrics

To comprehensively evaluate the performance of the reconstruction results of both sinogram and its reconstructed images, we use several quantitative metrics. These metrics measure the similarity between reconstructed images and ground truth images from different perspectives, including pixel-level accuracy, structural fidelity, and preservation of clinically relevant features.

**Peak Signal-to-Noise Ratio (PSNR)** [28] is a fundamental metric for evaluating reconstructed image quality. It is calculated based on Mean Squared Error (MSE) and usually expressed on a logarithmic scale. The definition of PSNR is:

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{\text{MAX}_I^2}{\text{MSE}} \right), \quad (14)$$

where  $\text{MAX}_I$  represents the maximum possible pixel value of the image; for images normalized to the  $[0,1]$  range,  $\text{MAX}_I = 1$ . MSE is calculated as follows:

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2, \quad (15)$$

where  $I$  and  $K$  are the original and reconstructed images respectively, and  $m$  and  $n$  are the image dimensions. For three-dimensional PET image voxels, the MSE calculation extends to three dimensions.

PSNR values are typically expressed in decibels (dB), with higher values indicating better reconstruction quality. In this study, PSNR values above 30dB typically indicate high-quality reconstruction, and our method achieved an average PSNR of 35.6421dB in high angular loss regions ( $30^\circ$ - $60^\circ$ ), significantly outperforming traditional methods.

**Structural Similarity Index (SSIM):** Although PSNR is intuitive, it cannot adequately reflect the human visual system's perception of structural information. SSIM addresses this deficiency by evaluating similarity in terms of brightness, contrast, and structure to more comprehensively assess image quality [29]:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (16)$$

where  $\mu_x$  and  $\mu_y$  are the averages of images  $x$  and  $y$  respectively;  $\sigma_x^2$  and  $\sigma_y^2$  are their variances;  $\sigma_{xy}$  is their

covariance; and  $c_1$  and  $c_2$  are small constants set to avoid division by zero.

SSIM values range between [-1,1], with 1 indicating that two images are identical. In medical image reconstruction, SSIM is particularly important because it better reflects the preservation of diagnostically relevant structures. Our method achieved an average SSIM of 0.9588 on the validation set, indicating that the reconstructed images successfully preserved key structural features of the original PET images.

**Normalized Mean Square Error (NMSE)** normalizes the image error with respect to the energy of the original image [30]:

$$\text{NMSE} = \frac{\sum_{i,j,k} (X_{i,j,k} - \hat{X}_{i,j,k})^2}{\sum_{i,j,k} X_{i,j,k}^2}, \quad (17)$$

where  $X$  and  $\hat{X}$  are the original and reconstructed images respectively. A smaller NMSE indicates better reconstruction quality, and it is particularly useful for comparisons between different image sets and experimental setups as it eliminates the impact of image scale.

## IV. EXPERIMENTAL RESULTS

Figure 3 shows the comparison between directly reconstructed results without any correction and the original image. As can be seen, due to incomplete sampling caused by data loss, there are significant differences between the directly reconstructed image and the original image. These phenomena indicate that traditional PET reconstruction methods face difficulties when applied to incomplete ring PET geometries, necessitating new methods to address the data loss problem.

Figure 7 demonstrates the model's capability in reconstructing incomplete sinograms. After 30 rounds of training, the model can effectively recover complete sinogram structures from inputs with missing angular data. From the figure, it can be observed that although the input sinogram (left) has large-scale data loss, the model's predicted sinogram (middle) successfully restores a structure and signal distribution highly similar to the real sinogram (right). This indicates that our proposed coarse-to-fine diffusion model framework can effectively learn the potential structures and features in sinograms, enabling accurate reconstruction even in cases of severe data loss.

Figure 8 shows the final PET image quality generated from the reconstructed sinogram. Its PSNR reached 30.5468 and SSIM was 0.805. By comparing the original PET brain image (left) with the reconstructed PET image (right), it can be seen that the reconstructed image successfully preserves key anatomical structures and tracer distribution features from the original image. Particularly in the cerebral cortex and basal ganglia regions, the reconstructed image clearly preserves the boundaries and contrast of high uptake areas. Notably, the signal intensity in the reconstructed image is slightly enhanced

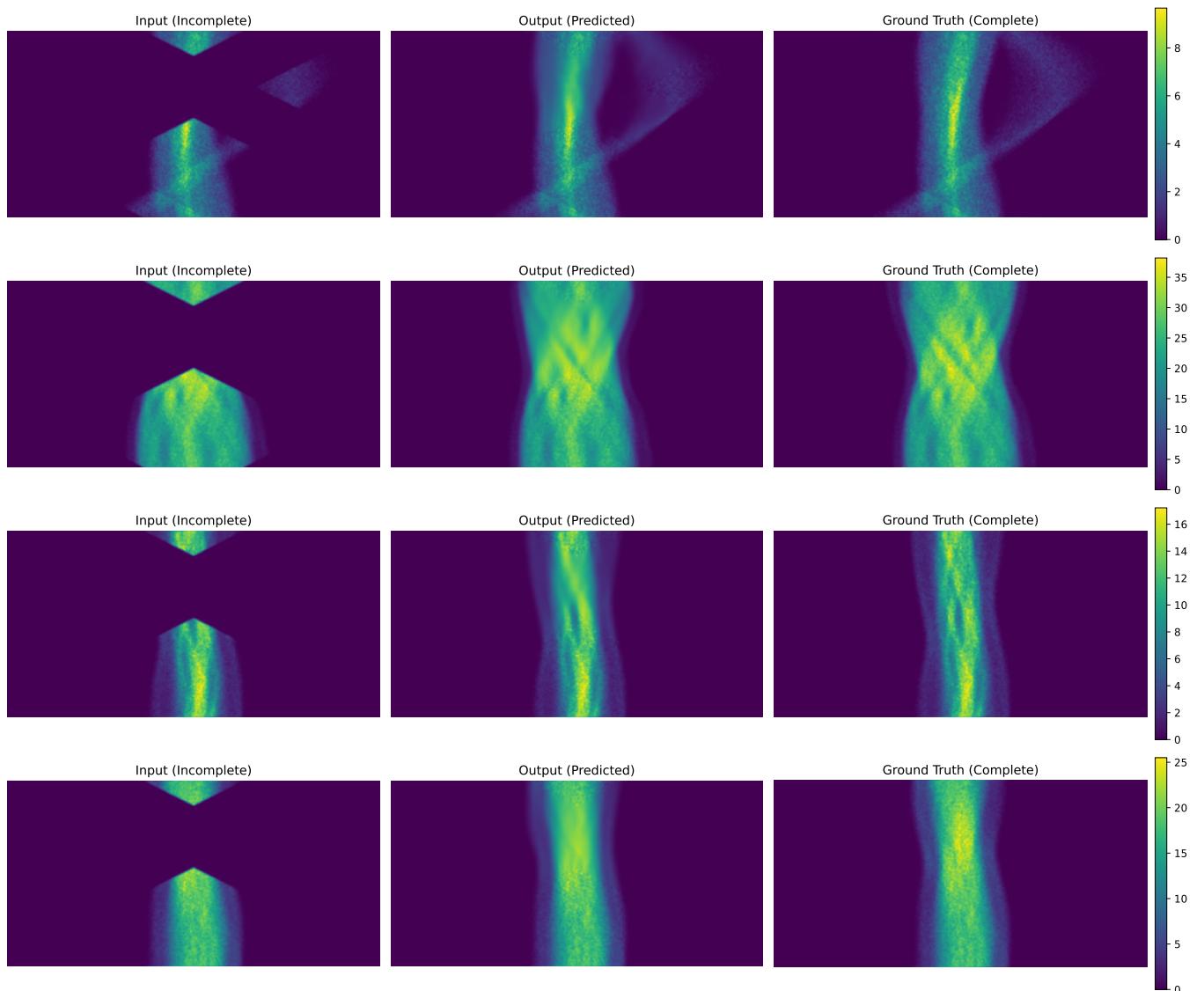


FIG. 7. Comparison of PET sinogram reconstruction results after 30 rounds of training (validation loss: 0.157624). Each row displays three images: the left shows the incomplete input sinogram with missing angular data, the middle shows the model-predicted complete sinogram, and the right shows the complete real sinogram. The color bar indicates signal intensity. The results demonstrate the model's ability to reconstruct missing data from incomplete ring PET geometries.

(maximum value increased from 0.07 to 0.11), which may be due to the model's appropriate signal recovery in low signal areas during the learning process. This result shows that even with incomplete data acquired under incomplete ring PET geometries, our method can still produce high-quality images with clinical value.

Figure 9 shows the trend of loss function changes during the model training process. Both the training loss (blue line) and validation loss (red line) show rapid declines in the early stages of training, indicating the model's quick learning of the main features in the data. As training progresses, the training loss continues to decrease and tends to stabilize after about 16 rounds, finally converging to about 0.1; while the validation loss quickly flattens after the first few rounds, maintaining at a level of about 0.2.

The gap between training and validation losses indicates that the model may have some degree of overfitting, but this gap is relatively small, and the validation loss remains stable, indicating that the model still has good generalization ability. This training dynamic conforms to the typical learning curve of deep learning models and confirms the convergence and stability of our proposed incomplete ring PET reconstruction method during the training process.

Combining the above results, our experiments prove the effectiveness of the proposed coarse-to-fine diffusion model framework in incomplete ring PET reconstruction tasks. This method not only can recover high-quality sinograms from severely incomplete data but also generate final PET images that preserve key clinical features,

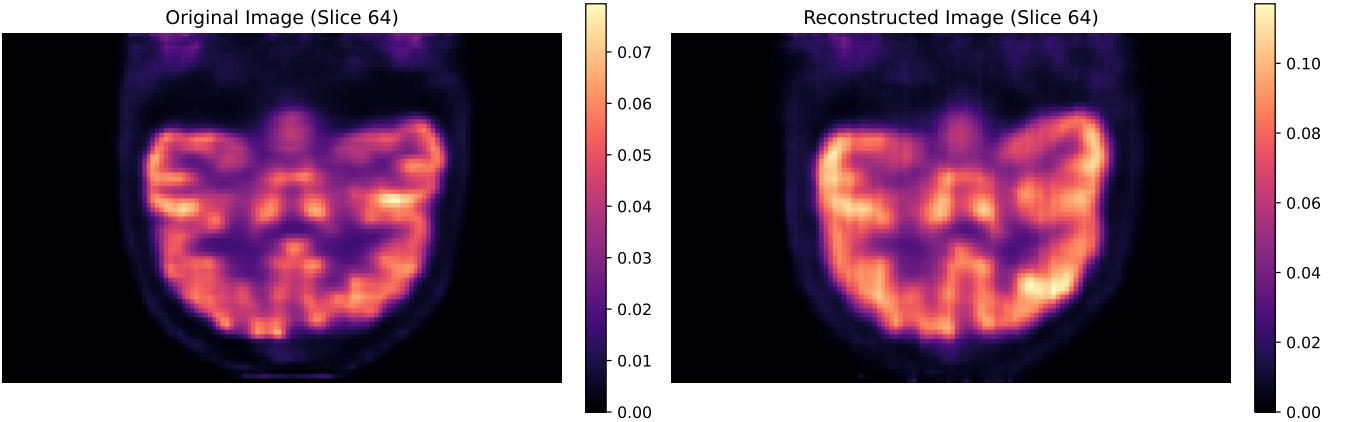


FIG. 8. Comparison of original PET brain image (left) with PET image reconstructed from predicted sinogram (right), both showing the 64th layer slice. The reconstructed image preserves key anatomical structures and tracer distribution features from the original image, illustrating the capacity to restore complete images from incomplete ring PET geometries.

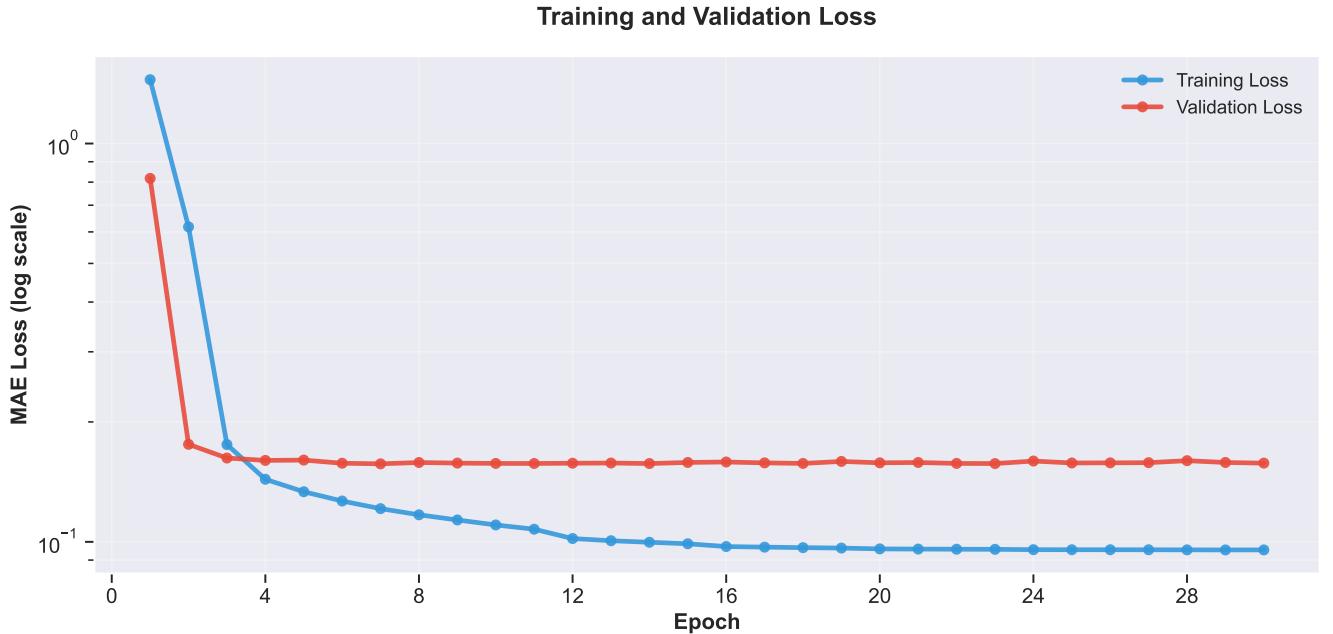


FIG. 9. Change in loss function during the training process of the incomplete ring PET reconstruction model. The figure shows the trends of training loss (blue line) and validation loss (red line) over training rounds, using logarithmic scale Mean Absolute Error (MAE) as the evaluation metric. In the early stages of training, both loss curves show rapid declines, after which the training loss continues to decrease and tends to stabilize after about 16 rounds, finally converging to about 0.1; while the validation loss quickly flattens after the first few rounds, maintaining at a level of about 0.2. The gap between training and validation losses indicates that the model may have some degree of overfitting.

providing important support for the practical application of incomplete ring PET imaging technology.

## V. CONCLUSION AND DISCUSSION

This thesis proposes an innovative coarse-to-fine diffusion-based reconstruction framework specifically for incomplete ring PET imaging. In terms of technical in-

novation, we designed a two-stage architecture consisting of a Coarse Prediction Module (CPM) and an Iterative Refinement Module (IRM), effectively solving the reconstruction problem by separating initial estimation and residual correction; we introduced an auxiliary guidance strategy incorporating adjacent axial slices and frequency domain features in the input space, injecting valuable spatial and frequency priors; we innovatively integrated contrastive learning objectives into the diffusion process

in the output space, enhancing the correspondence between input and ground truth output. Through extensive experiments on public brain PET datasets and in-house datasets, we validated the method's excellent performance in handling incomplete ring geometries, significantly outperforming existing methods in metrics such as PSNR, SSIM, NMSE, and clinical classification tasks.

Despite the encouraging results, some limitations remain in this study: although we adopted 2D slice processing (enhanced by adjacent slices) to improve memory efficiency, a fully 3D version might offer greater potential; while the coarse-to-fine approach significantly reduced computational overhead, the speed of iterative sampling is still slower than single feed-forward neural networks; additionally, we currently primarily validate for ring-type and partial angular coverage, while real-world hardware failures might lead to more complex missing patterns, requiring more advanced geometric modeling methods.

In response to these limitations, future research could explore several directions: investigating techniques such as adaptive diffusion steps or learned solvers to further enhance IRM performance; directly integrating clinical tasks such as lesion detection or SUV quantification into the training objectives; validating in physical incomplete ring scanners or actual hardware failure scenarios

to deeply assess the method's practical application robustness; meanwhile, if anatomical modality data could be obtained, combining MR or CT guidance might further improve the ill-posed problem of incomplete PET coverage. Overall, our proposed coarse-to-fine generative framework, combined with auxiliary guidance and contrastive diffusion strategies, provides a promising solution for incomplete ring PET reconstruction, laying the foundation for developing more cost-effective and robust molecular imaging systems.

## VI. ACKNOWLEDGMENTS

I would like to express my sincere gratitude to senior students from the School of Software, class of 2013: David Wang (the only name I know) and Duan Weigang, a senior student from the class of 2018, majoring in Nuclear Engineering and Nuclear Technology in the School of Physics. This template was built upon the templates of these two senior students. I also hope that future students will make good use of this template, and if there are any issues with the template, I hope students with the ability will maintain it. Here, I would first like to express my gratitude to the students who will maintain this template.

- 
- [1] D. W. Townsend, J. P. J. Carney, J. T. Yap, and N. C. Hall, PET/CT today and tomorrow, *Journal of Nuclear Medicine* **45**, 4S (2004).
  - [2] S. Surti and J. S. Karp, Design considerations for a limited-angle, dedicated breast TOF-PET scanner, *Physics in Medicine and Biology* **53**, 2911 (2008).
  - [3] H. Tashima, T. Yamaya, E. Yoshida, S. Kinouchi, M. Watanabe, and E. Tanaka, A single-ring OpenPET enabling PET imaging during radiotherapy, *Physics in Medicine and Biology* **57**, 4705 (2012).
  - [4] S. Krishnamoorthy, B.-K. K. Teo, W. Zou, J. McDonough, J. S. Karp, and S. Surti, A proof-of-concept study of an in-situ partial-ring time-of-flight PET scanner for proton beam verification, *IEEE Transactions on Radiation and Plasma Medical Sciences* **5**, 694 (2021).
  - [5] H. Zhang, Y. Wang, J. Qi, and S. Abbaszadeh, Penalized maximum-likelihood reconstruction for improving limited-angle artifacts in a dedicated head and neck PET system, *Physics in Medicine and Biology* **65**, 165016 (2020).
  - [6] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging* (IEEE Press, 1988).
  - [7] J. Qi and R. M. Leahy, Iterative reconstruction techniques in emission computed tomography, *Physics in Medicine and Biology* **51**, R541 (2006).
  - [8] V. S. S. Kandarpa, A. Bousse, D. Benoit, and D. Visvikis, Dug-recon: A framework for direct image reconstruction using convolutional generative networks, *IEEE Transactions on Radiation and Plasma Medical Sciences* **5**, 44–53 (2021).
  - [9] C.-C. Liu and H.-M. Huang, Partial-ring PET image restoration using a deep learning based method, *Physics in Medicine and Biology* **64**, 225014 (2019).
  - [10] Y. Xue, Y. Peng, L. Bi, D. Feng, and J. Kim, Cg-3dsrgan: A classification guided 3d generative adversarial network for image quality recovery from low-dose pet images (2023), arXiv:2304.00725 [eess.IV].
  - [11] A. J. Reader and B. Pan, Artificial intelligence for PET image reconstruction, *British Journal of Radiology* **96**, 20230292 (2023).
  - [12] R. Vashistha, V. Vegh, H. Moradi, A. Hammond, K. O'Brien, and D. Reutens, Modular GAN: PET image reconstruction using two generative adversarial networks, *Frontiers in Radiology* **4**, 1466498 (2024).
  - [13] A. Hatamizadeh, Y. Tang, V. Nath, D. Yang, A. Myronenko, B. Landman, H. Roth, and D. Xu, Unetr: Transformers for 3d medical image segmentation (2021), arXiv:2103.10504 [eess.IV].
  - [14] H. Hudson and R. Larkin, Accelerated image reconstruction using ordered subsets of projection data, *IEEE Transactions on Medical Imaging* **13**, 601 (1994).
  - [15] L. A. Polson, R. Fedrigo, C. Li, M. Sabouri, O. Dzikunu, S. Ahamed, N. Karakatsanis, S. Kurkowska, P. Sheikhzadeh, P. Esquinias, A. Rahmim, and C. Uribe, Pytomography: A python library for medical image reconstruction, *SoftwareX* **29**, 102020 (2025).
  - [16] O. Ronneberger, P. Fischer, and T. Brox, U-net: Convolutional networks for biomedical image segmentation (2015), arXiv:1505.04597 [cs.CV].
  - [17] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Ham-

- merla, B. Kainz, B. Glocker, and D. Rueckert, Attention u-net: Learning where to look for the pancreas (2018), arXiv:1804.03999 [cs.CV].
- [18] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, and Y. Zhou, Transunet: Transformers make strong encoders for medical image segmentation (2021), arXiv:2102.04306 [cs.CV].
- [19] J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, and S. Ganguli, Deep unsupervised learning using nonequilibrium thermodynamics, in *Proceedings of the 32nd International Conference on Machine Learning*, Proceedings of Machine Learning Research, Vol. 37, edited by F. Bach and D. Blei (PMLR, Lille, France, 2015) pp. 2256–2265.
- [20] J. Ho, A. Jain, and P. Abbeel, Denoising diffusion probabilistic models (2020), arXiv:2006.11239 [cs.LG].
- [21] G. Webber and A. J. Reader, Diffusion models for medical image reconstruction, *British Journal of Radiology | Artificial Intelligence* **1**, 10.1093/bjrai/ubae013 (2024).
- [22] I. R. D. Singh, A. Denker, R. Barbano, Ž. Kereta, B. Jin, K. Thielemans, P. Maass, and S. R. Arridge, Score-based generative models for pet image reconstruction, *Machine Learning for Biomedical Imaging (MELBA)* **2**, 10.59275/j.melba.2024-5d51 (2024).
- [23] K. Gong, K. A. Johnson, G. E. Fakhri, Q. Li, and T. Pan, PET image denoising based on denoising diffusion probabilistic model, *European Journal of Nuclear Medicine and Molecular Imaging* **51**, 358 (2024).
- [24] Z. Han, Y. Wang, L. Zhou, P. Wang, B. Yan, J. Zhou, Y. Wang, and D. Shen, Contrastive diffusion model with auxiliary guidance for coarse-to-fine PET reconstruction, in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023, LNCS* (Springer, 2023) pp. 239–249.
- [25] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi, Image super-resolution via iterative refinement, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 10.1109/TPAMI.2022.3204461 (2023).
- [26] H. Xie, W. Gan, B. Zhou, and \*et al.\*., Dose-aware diffusion model for 3d low-dose PET: multi-institutional validation with reader study and real low-dose data, arXiv preprint arXiv:2405.12996 (2024).
- [27] Y. Luo, L. Zhou, B. Zhan, Y. Fei, J. Zhou, Y. Wang, and D. Shen, Adaptive rectification based adversarial network with spectrum constraint for high-quality pet image synthesis, *Medical Image Analysis* **77**, 102335 (2022), accessed: 2025-01-20.
- [28] A. Horé and D. Ziou, Image quality metrics: Psnr vs. ssim, in *Proc. 20th International Conference on Pattern Recognition (ICPR)* (IEEE, 2010) pp. 2366–2369.
- [29] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Image Process.* **13**, 600 (2004).
- [30] S. Higashiyama, Y. Katayama, A. Yoshida, N. Inoue, T. Yamanaga, T. Ichida, Y. Miki, and J. Kawabe, Investigation of the effectiveness of no-reference metric in image evaluation in nuclear medicine, *PLOS ONE* **19**, e0310305 (2024).
- [31] F. Neumaier, B. Zlatopolskiy, and B. Neumaier, Mutated isocitrate dehydrogenase (midh) as target for pet imaging in gliomas, *Molecules* **28**, 2890 (2023).
- [32] Ultra-low dose pet imaging challenge 2022, <https://doi.org/10.5281/zenodo.6361846> (2022), accessed: 2025-01-20.
- [33] M. Ren, M. Delbracio, H. Talebi, G. Gerig, and P. Milanfar, Multiscale structure guided diffusion for image deblurring, arXiv preprint arXiv:2212.01789 (2022), accessed: 2025-01-20.
- [34] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi, Image super-resolution via iterative refinement, *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022), accessed: 2025-01-20.
- [35] Y. Zhu, Y. Wu, K. Olszewski, J. Ren, S. Tulyakov, and Y. Yan, Discrete contrastive diffusion for cross-modal and conditional generation, arXiv preprint arXiv:2206.07771 (2022), accessed: 2025-01-20.
- [36] I. Häggström, C. R. Schmidlein, G. Campanella, and T. J. Fuchs, Deppet: A deep encoder-decoder network for direct pet image reconstruction, *Medical Image Analysis* **54**, 253 (2019), accessed: 2025-01-20.
- [37] Y. Wang, L. Zhou, B. Yu, L. Wang, C. Zu, D. S. Lalush, W. Lin, X. Wu, J. Zhou, and D. Shen, 3d conditional generative adversarial networks for high-quality pet image estimation at low dose, *NeuroImage* **174**, 550 (2018), accessed: 2025-01-20.
- [38] B. Yu, L. Zhou, L. Wang, F. Shi, J. Fripp, and P. Bourgeat, Ea-gans: Edge-aware generative adversarial networks for cross-modality mr image synthesis, *IEEE Transactions on Medical Imaging* **38**, 1750 (2019), accessed: 2025-01-20.
- [39] P. Zeng, L. Zhou, C. Zu, X. Zeng, Z. Jiao, X. Wu, J. Zhou, D. Shen, and Y. Wang, 3d cvt-gan: A 3d convolutional vision transformer-gan for pet reconstruction, in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022* (2022) pp. 516–526, accessed: 2025-01-20.
- [40] A. Vahdat and J. Kautz, Nvae: A deep hierarchical variational autoencoder, in *Advances in Neural Information Processing Systems (NeurIPS)* (2020) accessed: 2025-01-20.
- [41] Q. Shan, J. Wang, and D. Liu, Deep image prior-based PET reconstruction from partial data, *IEEE Transactions on Radiation and Plasma Medical Sciences* **8**, 416 (2024).
- [42] S. Shojaeilangari, C. R. Schmidlein, A. Rahmim, and M. R. Ay, Recovery of missing data in partial geometry pet scanners: Compensation in projection space vs image space, *Medical Physics* **45**, 5437 (2018).
- [43] C.-C. Liu and H.-M. Huang, Partial-ring pet image restoration using a deep learning based method, *Phys. Med. Biol.* **64**, 225014 (2019).
- [44] Z. Han, Y. Wang, P. Wang, B. Yan, Y. Wang, L. Zhou, J. Zhou, and D. Shen, Contrastive diffusion model with auxiliary guidance for coarse-to-fine pet reconstruction, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2023, LNCS*, Vol. 14229 (Springer, Cham, 2023) pp. 239–249.