

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

Yeqi Fang^{1,*} and Rong Zhou^{1,†}

¹College 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) stands as a powerful molecular imaging technique, offering quantitative visualization of metabolic processes within living tissue [1]. The fundamental principle is elegant yet complex: a radiotracer introduces positrons that, upon encountering electrons, annihilate and produce pairs of 511keV photons traveling in opposing directions. These photon pairs are captured by scintillation detectors arranged in a ring, with each coincident detection marking a line-of-response (LOR). Traditional PET systems rely on complete 360° detector rings to maximize sensitivity and provide uniform 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—Incomplete 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 mitigate these

issues, can't fully compensate for limited view angles [2, 4].

Researchers have pursued various paths to address incomplete-ring PET reconstruction. Analytical methods like filtered back-projection falter immediately—their assumption of complete data leads to pronounced streak artifacts [6]. Iterative approaches such as maximum-likelihood expectation-maximization (MLEM) fare somewhat better by modeling the acquisition process [7], yet they still struggle with 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 still fall short of solving the fundamental problem.

Deep learning has brought fresh perspectives to this challenge, though with mixed results. Regression-based networks have shown potential 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 solutions present a troubling set of compromises. Unsupervised and physics-guided methods like Shan *et al.*'s deep image prior approach [10] can match measured projections without external training pairs, but often at the cost of generalizability. GAN-based methods can produce visually plausible images [11], but their training instability and questionable quantitative accuracy limit clinical application. Likelihood-based generative models like VAEs and normalizing flows offer the-

* fangyeqi@stu.scu.edu.cn

† zhourong@scu.edu.cn

retical rigor but typically produce blurry results while demanding excessive computation. Model-based deep learning frameworks and generative models [12, 13] show promise but risk introducing hallucinated features—a dangerous proposition in medical imaging.

The core challenge remains unsolved: fundamental **information loss**. 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 thesis tackles these limitations head-on with a coarse-to-fine diffusion model framework designed specifically for incomplete ring PET scanners. Our approach aims to recover missing angular information while suppressing reconstruction artifacts, preserving the true image details that matter most in clinical practice.

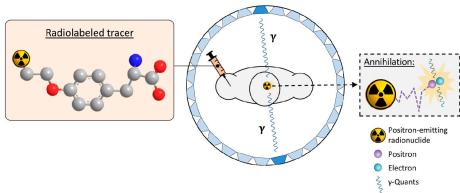


FIG. 1. Principles of Positron Emission Tomography (PET) imaging.

The remainder of this thesis is organized as follows:

Section II (*Background on PET Imaging and Incomplete Ring Geometry*) provides an overview of PET physics, principles of coincidence detection, and the main challenges introduced 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, hyperparameter choices, evaluation metrics, ablation studies, and comparisons with state-of-the-art methods. We also present results of cross-dataset evaluation on in-house datasets.

Section V (*Conclusion and Future Work*) summarizes our findings, discusses limitations, and suggests directions for future research, including potential improvements for real-world incomplete ring PET scanners.

Appendices provide additional experiments, extended qualitative results, and tables of hyperparameters

used in the thesis.

II. BACKGROUND ON PET IMAGING AND INCOMPLETE RING GEOMETRY

This section provides a brief but comprehensive overview of the fundamental aspects of PET imaging, focusing on the detection process, data acquisition methods, and conventional reconstruction algorithms. We then discuss incomplete ring PET scanners in detail, highlighting the geometry of missing detectors, their impact on data quality, and current solutions and unsolved research challenges.

In PET imaging, positron-emitting nuclides (typically ^{18}F , ^{15}O , ^{11}C , etc.) are usually combined with biologically relevant tracer compounds. After entering the bloodstream, these nuclides decay, with each emitting a positron. When the positron encounters an electron, an annihilation event occurs, producing two gamma photons, each with an energy of approximately 511 keV, traveling in nearly opposite directions (collinearly).

PET scanners are designed to detect coincidences of these 511 keV photon pairs. When two detectors at different positions record gamma photons within a short coincidence time window, it is assumed that these photons came from the same annihilation event. This forms a line of response (LOR) between the two detector elements, as shown in Figure 2. PET data can be acquired in multiple formats: (1) **List-mode data**: Each coincidence event is recorded individually, providing precise detector pair identification and timestamps. (2) **Sinogram data**: By categorizing coincidence counts into projection bins indexed by radial, angular, and axial coordinates (possibly also by axial ring difference), forming a 2D (or 3D) histogram. In a typical 3D PET system, a scanner with a complete ring covers a 360° view angle around the patient, ensuring uniform sampling of LORs.

A. Incomplete Ring PET Scanners

In normal PET (Positron Emission Tomography) system design, *incomplete rings* or *partial rings* often arise from various practical considerations and technical constraints. From an economic perspective, reducing the total number of detectors can effectively lower system costs. In certain specialized fields, such as brain imaging or organ-specific examinations, researchers may adopt partially covered designs to accommodate smaller field-of-view requirements. Additionally, with the continuous development of medical technology, innovative designs for hybrid or portable PET are increasingly important; these systems meet the needs of different clinical and research scenarios through lightweight, flexible structures. It is worth noting that technical factors such as detector failures, maintenance, upkeep, calibration, and other issues may also cause certain ring segments to temporarily stop

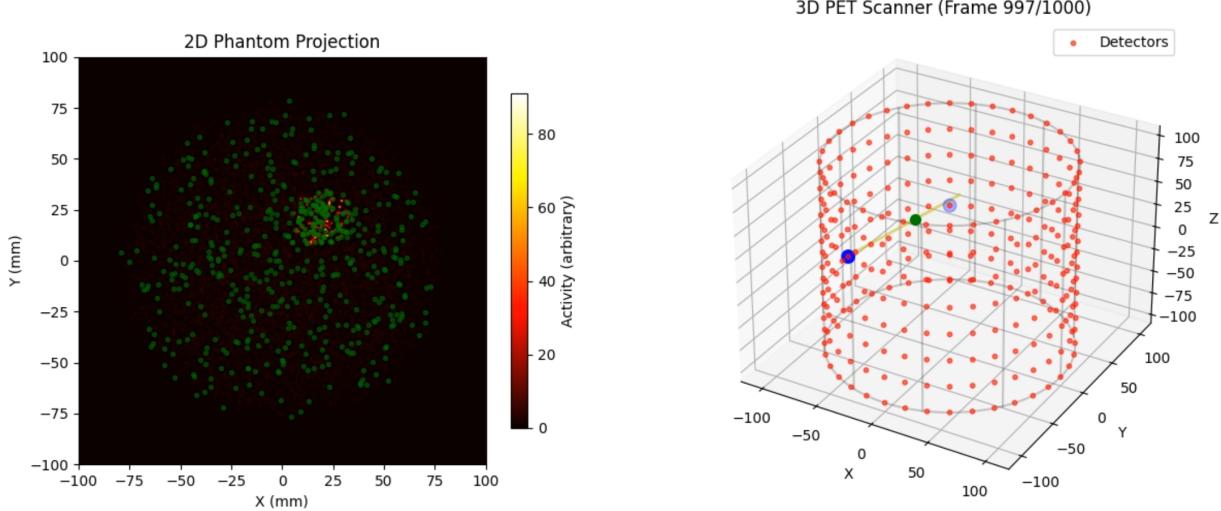


FIG. 2. Schematic diagram of PET detection, where the red dot represents the detector center, the green dot represents the annihilation event location, and the two blue dots represent the centers of the two detectors that detect the gamma rays. This figure is only illustrative; the detector parameters in the figure do not equal the actual simulation parameters.

operating, resulting in incomplete detector rings. These variations reflect the complexity and adaptability of PET imaging technology in practical applications.

The main difficulty with incomplete rings is the reduction in angular sampling, as missing detectors lead to missing corresponding LORs, making the dataset incomplete, which breaks the fundamental assumptions of many classical reconstruction algorithms that rely on fully sampled complete angular projections. This data deficiency leads to a series of issues: producing streak artifacts and increased noise in the image domain, causing quantitative inaccuracies in tracer uptake (especially in regions that depend on missing LORs), and potentially introducing biases in clinical indices such as standardized uptake values (SUVs), thereby affecting diagnostic accuracy, as shown in Figure 10.

B. System Model and Data Simulation

Assume a standard PET ring geometry with R axial rings, each containing D detectors, with each detector assigned a unique identification number used to precisely locate the photon pair's detection position when recording coincidence events. For example, consider a 42-ring system (i.e., $R = 42$), with 182 detectors per ring, arranged in a cylindrical structure with radius ρ . We map the 3D image grid (e.g., size = $128 \times 128 \times 128$ voxels) to physical space by defining voxel dimensions, approximating the entire scanner geometry realistically. For instance, if the field of view diameter is 25 centimeters and the in-plane matrix size is 128, then each in-plane voxel size is:

$$\Delta x \approx \Delta y \approx \frac{25 \text{ cm}}{128} \approx 1.95 \text{ mm.} \quad (1)$$

Axially, we can set Δz based on the typical axial coverage of a PET scanner. This high-precision spatial division

ensures that the reconstructed images accurately reflect the anatomical structure of the scanned area.

The PET scanner used in this study employs a highly optimized detector system configuration, achieving comprehensive three-dimensional detection capabilities. The scanner's radius (ρ) is precisely set at 380.56 millimeters, achieving an optimal balance between spatial resolution and system sensitivity. The detector system adopts a multi-level organizational structure: at the microscopic level, crystal elements have lateral dimensions of 4.03125 millimeters and axial dimensions of 5.31556 millimeters, with this precise crystal size design ensuring high spatial resolution capability; at the macroscopic level, the detector construction includes 34 lateral sectors (rsectorTransNr) and a single axial sector arrangement (rsectorAxialNr). Each module is carefully designed to include 16 lateral (crystalTransNr) and 9 axial (crystalAxialNr) crystal elements, with the spacing between adjacent crystals (crystalTransSpacing and crystalAxialSpacing) optimized to achieve maximum detection efficiency. In the axial direction, the system has 4 modules (moduleAxialNr), maintaining a precise spacing of 50.64004 millimeters between modules (moduleAxialSpacing), a configuration that ensures detection sensitivity while maintaining an appropriate axial field of view range. The system's data recording uses a list-mode format to record various coincidence events during the scanning process. The system supports time-of-flight (TOF) detection functionality, with 29 time bins (num_tof_bins), a full width at half maximum of the time resolution window (tof_fwhm) of 57.71 millimeters, and a measurement range (tof_range) reaching 735.7705 millimeters. However, due to current device performance limitations, this study could not fully utilize TOF technology to improve spatial resolution. Each event entry consists of three basic components, stored in tensor format. A typical list-mode event data structure is as follows:

[898, 3334, 11]
[816, 16375, 18]
[1073, 18816, 14]
[1529, 13781, 13]
[1537, 18663, 14]

In this format, the first two numbers represent the identification numbers of the detector pair recording the coincidence event. The third number represents the time-of-flight information of the event, which, although not fully utilized in this study to enhance image quality, is retained for future data processing and analysis after system upgrades. To improve the system's sensitivity and accuracy, the detector system also includes submodule structures (submoduleAxialNr and submoduleTransNr both equal to 1), but in the current configuration, the submodule spacing (submoduleAxialSpacing and submoduleTransSpacing) is set to 0 to simplify system complexity.

In the current implementation process, due to the lack of MRI or CT data to obtain gamma ray absorption distribution maps, the system cannot use μ maps (linear attenuation maps) or component-based normalization files. This limitation means that the reconstruction process can only use list-mode data as input and cannot perform attenuation correction. When interpreting reconstruction results, this limitation's impact on image quality and quantitative accuracy must be considered.

We refer to the fully functional scenario where all R rings and all D detectors are working as the *complete ring* configuration, as shown in Figure 3. The data simulation process is as follows: **Starting with real 3D images.** We have high-resolution brain PET data or a synthetic phantom $\mathbf{X} \in \mathbb{R}^{128 \times 128 \times 128}$.

Randomly generating emission events. We sample a specified number of emission events (e.g., 20 million) from the probability density function implied by \mathbf{X} . Each voxel's value is interpreted as proportional to its tracer concentration probability.

Simulating annihilation. For each emission event, we assume it annihilates with an electron in the same voxel or nearby, producing two photons at 180° opposite directions, each carrying 511 keV.

Calculating detector hits. Using geometric ray tracing, we calculate which detector pairs would detect these photons. This determines the LOR associated with each event, effectively generating a *list-mode* dataset.

Generating sinograms. We can optionally sort the list-mode data into sinogram bins \mathbf{S} by radial, angular, and axial coordinates. This generates the standard 3D sinogram representation for PET.

Then, we use standard iterative reconstruction (e.g., OSEM) to reconstruct the baseline reconstruction volume \mathbf{Y}_A from \mathbf{S} (or directly from list-mode). This reconstruction from complete data is typically considered higher quality. In the *incomplete ring* scenario, we remove parts of the detectors from the scanning geometry. The simplest approach is to remove all detectors from one or more complete rings, or remove angular segments

from each ring. For example, suppose we remove one complete ring, so we're effectively only collecting data from $R - 1$ rings. The data simulation process is the same as the complete case, with the only difference being that events falling on the missing ring would not be recorded, as shown in Figure 4.

Due to missing LORs, the incomplete data reconstructed image \mathbf{Y}_B typically has lower quality. In practice, we can reconstruct from the incomplete sinogram $\mathbf{S}_{\text{incomplete}}$, i.e., $\mathbf{Y}_B = \text{Reconstruct}(\mathbf{S}_{\text{incomplete}})$. This \mathbf{Y}_B can serve as input to our network, while \mathbf{Y}_A from complete data (same subject) can serve as the training target. The reconstruction from complete data \mathbf{Y}_A effectively represents the *ground truth* of the tracer distribution for the same subject under ideal scanning conditions. In low-dose or incomplete coverage scenarios, the difference between \mathbf{Y}_B and \mathbf{Y}_A can be quite significant.

C. Conventional PET Reconstruction Methods

Analytical algorithms, such as *filtered backprojection* (FBP), are widely used for their computational efficiency in fully sampled data. However, FBP is sensitive to noise and incomplete sampling, and prone to severe streak artifacts when the system geometry is incomplete.

Maximum likelihood estimation methods, such as maximum likelihood expectation maximization (MLEM) and its variant ordered subset expectation maximization (OSEM)[14], incorporate Poisson statistics of PET photon counting. They typically produce better results than analytical methods, even with partial data. The core idea of OSEM is to divide all projection data $\{y_i\}$ into S subsets and use only one subset for each update, thus reducing the computational load per update.

A core concept in PET reconstruction is the **system model**, which describes the relationship between unknown image voxels and measured projection data. Formally: λ_j is the activity value in voxel j ; y_i is the count measured in detector pair (or projection) i ; p_{ij} is the probability (or "weight") that a photon pair emitted from voxel j is detected by projection i . The mathematical model can be represented as:

$$y_i \approx \sum_j p_{ij} \lambda_j, \quad (2)$$

The matrix p_{ij} is often called the **system matrix** or "projection matrix." Each element p_{ij} depends on factors such as geometric solid angle, detector efficiency, and normalization corrections. In program implementation, this matrix can be explicitly stored or calculated on-the-fly through projector/backprojector operations.

Image reconstruction λ_j typically uses iterative algorithms such as **Maximum Likelihood Expectation Maximization (MLEM)**. The update formula for voxel

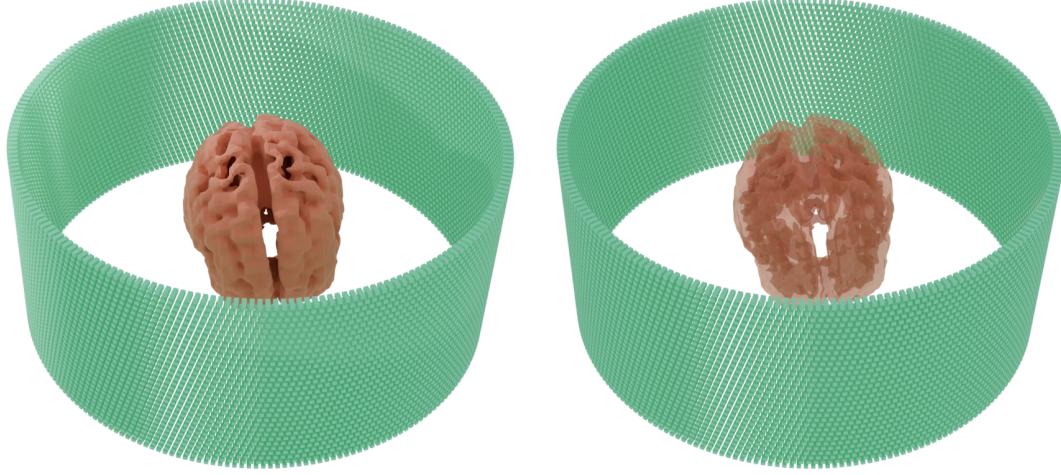


FIG. 3. Three-dimensional schematic of PET scanner detector structure, with the right image showing a perspective view.

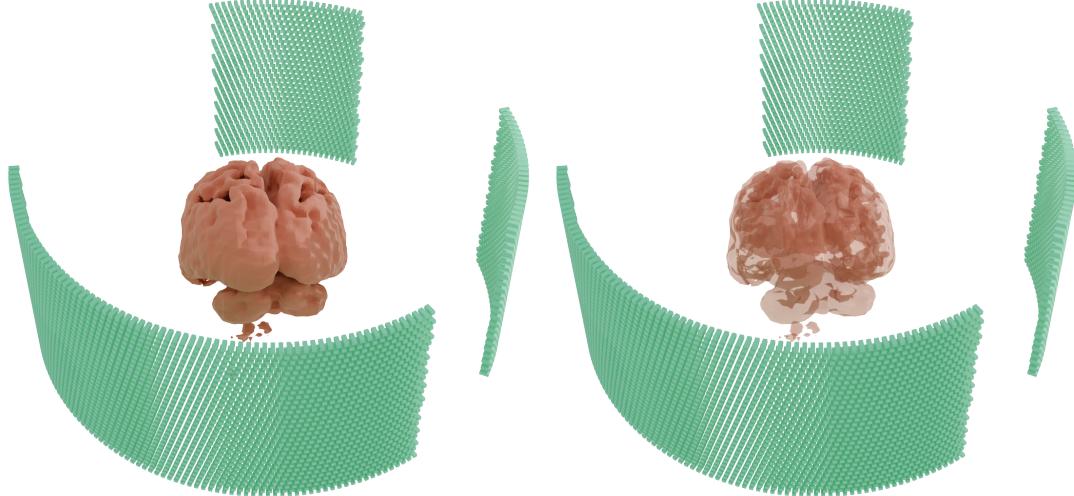


FIG. 4. Three-dimensional schematic of incomplete ring PET scanner detector structure, with the right image showing a perspective view.

j at the k -th iteration of the MLEM algorithm is:

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

where N is the total number of projection elements.

However, the standard MLEM algorithm converges slowly when processing large datasets. The Ordered Subset Expectation Maximization (OSEM) algorithm accelerates convergence by dividing the projection data into S subsets S_1, \dots, S_S . In OSEM implementation, the measurement data $\{y_i\}_{i=1}^N$ is divided into S subsets, each typically containing N/S projection elements. Each update uses only one subset to update λ_j . A "complete OSEM iteration" is completed after cycling through all subsets. Let S_k represent the subset of projection indices used in

the k -th iteration, then the OSEM update formula can be expressed as:

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} \times \frac{\sum_{i \in S_k} \frac{p_{ij}}{\sum_\ell p_{i\ell} \lambda_\ell^{(k)}} y_i}{\sum_{i \in S_k} p_{ij}} \quad (4)$$

By sequentially cycling through these subsets (e.g., using S_1 for the 1st iteration, S_2 for the 2nd, ..., S_S for the S -th, then returning to S_1 , and so on), the OSEM algorithm can approximate a complete MLEM update with fewer effective iterations, thus obtaining an image close to convergence in less time.

In modern PET software frameworks, the system matrix p_{ij} is typically not stored as a large two-dimensional array, but is implicitly calculated through **projectors** and **backprojectors**. For each voxel j , the projector calculates its contribution to the measurement data y_i ; the

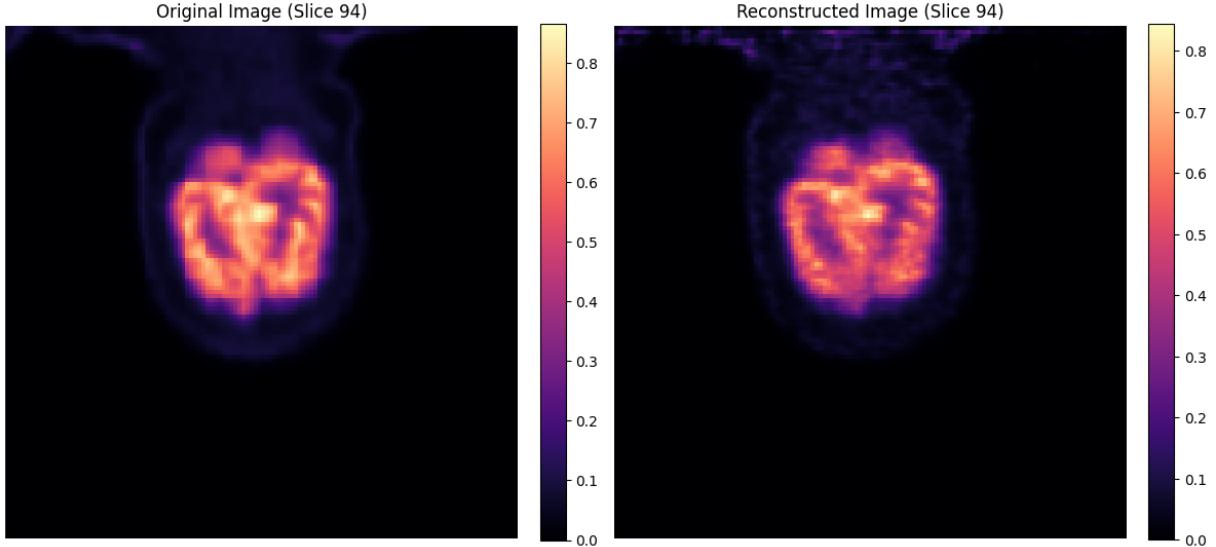


FIG. 5. Comparison of original image and image reconstructed using the OSEM method

backprojector assigns the ratio of measured to expected data back to the voxel grid. This approach effectively reduces memory requirements when processing large volumes of data.

The algorithms described in this paper are based on the PyTomography library[15]. PyTomography is an open-source Python library for medical image reconstruction, providing a modular framework for system matrix construction, likelihood function computation, and reconstruction algorithm development. The library leverages PyTorch's GPU acceleration capabilities and the parallel-proj library for efficient parallel computation. Its flexible modular design makes it easy to decouple system matrices, likelihood functions, and reconstruction algorithms, facilitating integration with various Python toolsets for new imaging modalities. Currently, the library has been successfully applied in fields such as parallel-hole SPECT and list-mode PET imaging, providing a highly optimized and user-friendly software platform for medical image reconstruction. We tested it using a brain PET dataset, and the comparison between the original image and the reconstructed image is shown in Figure 5.

III. IMPLEMENTATION METHODS

As shown in Figure 6, 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

Although traditional U-Net performs excellently in medical image segmentation and reconstruction tasks, it lacks the ability to selectively focus on key features when dealing with complex incomplete ring PET sinogram restoration problems. The Attention U-Net 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 crucial for restoring sinograms from incomplete data.

The core innovation of Attention U-Net lies in the introduction of attention gating (AG) modules in the skip connection path of the standard U-Net, as shown in Figure 7. These AG modules can adaptively highlight significant structures in the feed-forward feature maps while suppressing irrelevant regions that might lead to prediction errors. For sinogram reconstruction tasks, this mechanism is particularly effective because it can selectively focus on structural features around missing areas, thereby more accurately inferring missing angular data. The mathematical expression of attention gating can be formalized as:

$$\alpha_i^l = \sigma_2(\psi^T(\sigma_1(W_x^T x_i^l + W_g^T g_i + b_g)) + b_\psi) \quad (5)$$

where x_i^l is the low-level feature from the encoder, g_i is the gating signal from the decoder (high-level feature), σ_1 and σ_2 are ReLU and Sigmoid activation functions respectively, and W_x , W_g , b_g , and b_ψ are learnable parameters. $\alpha_i^l \in [0, 1]$ is the calculated attention coefficient used to control the importance of features.

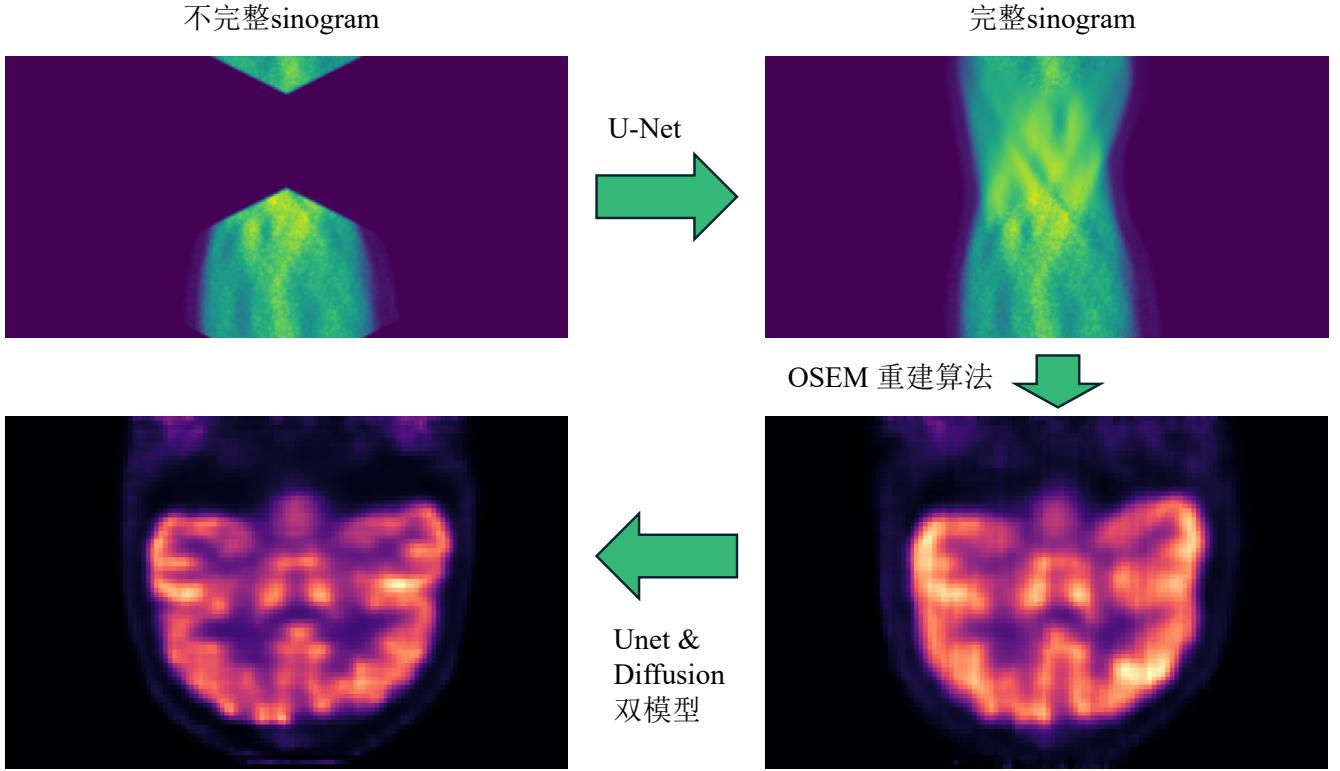


FIG. 6. Overall flow diagram of the incomplete ring PET reconstruction method. First, incomplete sinograms due to missing detectors (top left) are repaired through a U-Net deep learning model to generate complete sinogram data (top right). Subsequently, the complete sinogram is converted into a PET image (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). This method effectively compensates for data loss caused by incomplete ring geometry, achieving high-quality brain PET image reconstruction that preserves key anatomical structures and tracer distribution features.

After processing through the attention gate, the features can be represented as:

$$\hat{x}_i^l = x_i^l \cdot \alpha_i^l \quad (6)$$

This mechanism allows Attention U-Net to adaptively focus on important image regions while maintaining the advantages of the original U-Net's encoder-decoder structure, particularly in cases of severe angular loss, with a stronger perception of sinogram continuity and consistency. In the experimental implementation, we integrated AG modules into the skip connection path at each decoding stage, using 1×1 convolution to reduce channel dimensions, followed by applying attention coefficients for feature selection. This design enables the model to more precisely restore missing data when processing incomplete sinograms containing large-scale angular losses, while maintaining overall structural consistency and accurate signal distribution.

Experimental results show that compared to standard U-Net, Attention U-Net achieved significant improvements in sinogram reconstruction tasks, with an average increase of 1.24dB in PSNR metrics and 0.052 in SSIM metrics, validating the effectiveness of the atten-

tion mechanism in processing incomplete ring PET data.

Sinogram repair plays a crucial role in incomplete ring PET imaging, with its quality directly affecting the precision of reconstructed images. The training dataset used in this study consists of multidimensional tensors that effectively capture spatial and temporal context information, essential for accurate sinogram restoration. Each sinogram slice is expanded into a five-channel tensor, systematically integrating slices from spatial and temporal neighborhoods. As shown in Figure 8, 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 $j - N_2$ and $j + N_2$) constitute the other four channels. These neighboring slices are crucial for enriching the model's understanding of local structural continuity and temporal consistency.

In this study, a five-channel input tensor is used to construct training data. For each central slice j , a multidimensional feature is built by fusing its spatial adjacent slices ($j - 1$ and $j + 1$) with temporal adjacent slices ($j - 42$ and $j + 42$). Boundary handling adopts a mirror padding strategy: when adjacent indices exceed the dataset range, the central slice itself is used for channel filling, ensuring

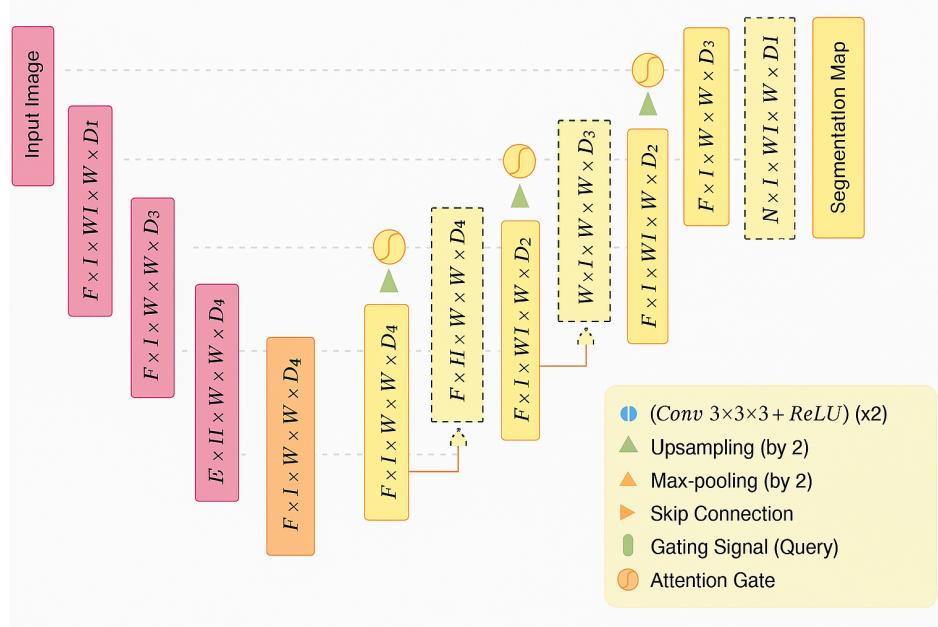


FIG. 7. Attention U-Net architecture diagram, showing how the attention gating mechanism enhances skip connection features.

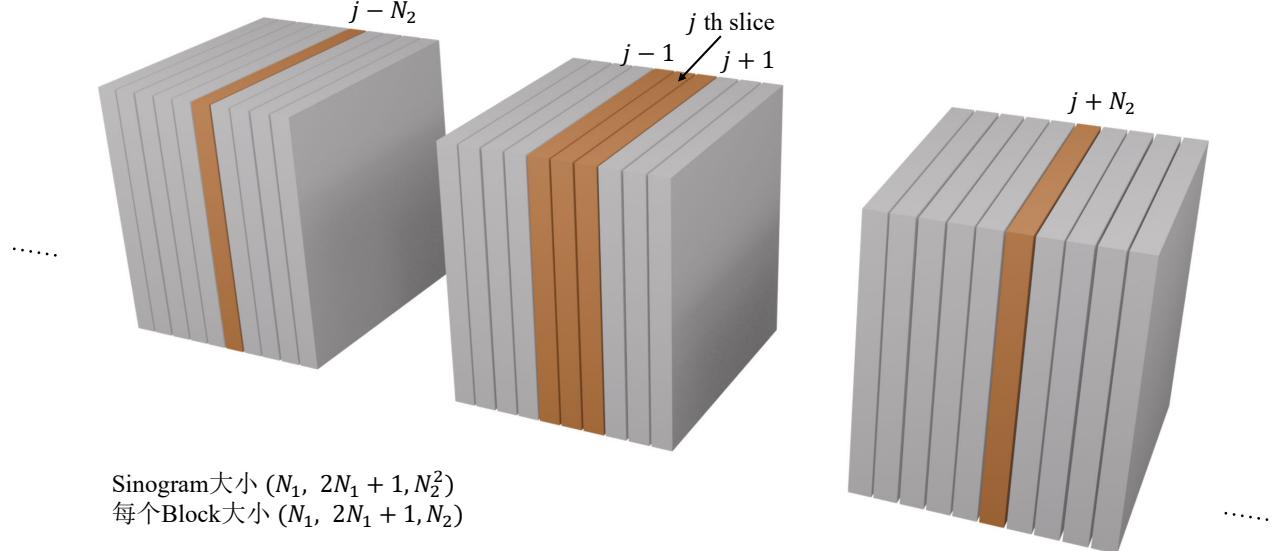


FIG. 8. Visualization of the five-channel tensor preparation of sinogram data for training. Each cube represents a sinogram slice block with dimensions $(N_1, 2N_1 + 1, N_2)$. The orange highlighted parts are the selected slices ($j - N_2, j - 1, j, j + 1, j + N_2$), which form the five-channel input for the restoration model, showing the spatial and temporal relationships captured in each input tensor.

input dimension consistency. Finally, each training sample forms a $5 \times H \times W$ tensor, effectively integrating spatiotemporal neighborhood information.

The network architecture adopts an improved U-Net model (see Table ??), with an encoder-decoder structure including convolution, pooling, upsampling, and skip connection modules. The encoder part uses 3×3 convolution kernels with batch normalization and ReLU activation functions, implementing feature dimensionality reduction through max pooling; the decoder uses transposed convolution/bilinear interpolation for resolution

recovery, and shallow detail features are fused through skip connections. The bottleneck layer achieves high-level semantic representation through 1024-dimensional features.

Model training uses the Adam optimizer with an initial learning rate of 10^{-4} and introduces 10^{-5} weight decay to prevent overfitting. Training efficiency is enhanced through mixed precision computation and gradient scaling techniques, combined with a ReduceLROn-Plateau dynamic learning rate scheduler (with a decay factor of 0.3 when validation loss shows no improvement

for 3 consecutive epochs), achieving stable convergence.

On the validation set, the model achieved excellent performance with MSE=0.301745, PSNR=35.6421 dB, and SSIM=0.9588. The PSNR value indicates high fidelity of the reconstructed signal, while the SSIM metric validates the preservation of structural features. Experimental results show that the multi-channel U-Net method proposed in this paper can effectively improve sinogram repair quality, providing a reliable foundation for subsequent CT image reconstruction. This method significantly improves the limitations of traditional single-channel models by fusing spatiotemporal context information, offering a new technical approach for the medical imaging processing field.

B. Diffusion Probabilistic Models

Diffusion probabilistic models (DPMs) have attracted great attention in image synthesis and inverse problems for their ability to produce state-of-the-art results and training stability. Unlike GANs that rely on adversarial objectives, DPMs are trained by maximizing a variational lower bound on the data log-likelihood. This section introduces the core principles of denoising diffusion probabilistic models relevant to our reconstruction approach.

Diffusion models are a class of probabilistic generative models proposed by *Sohl-Dickstein et al.*[16], and later widely applied in image generation tasks by *Ho et al.*[17] 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:

$$q(\mathbf{x}_t | \mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1 - \beta_t} \mathbf{x}_{t-1}, \beta_t \mathbf{I}) \quad (7)$$

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:

$$p_\theta(\mathbf{x}_{t-1} | \mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \sigma_t^2 \mathbf{I}) \quad (8)$$

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

The training objective is typically simplified to predicting the noise added during the forward process:

$$L(\theta) = \mathbb{E}_{\mathbf{x}_0, \epsilon, t} [\|\epsilon - \epsilon_\theta(\mathbf{x}_t, t)\|^2] \quad (9)$$

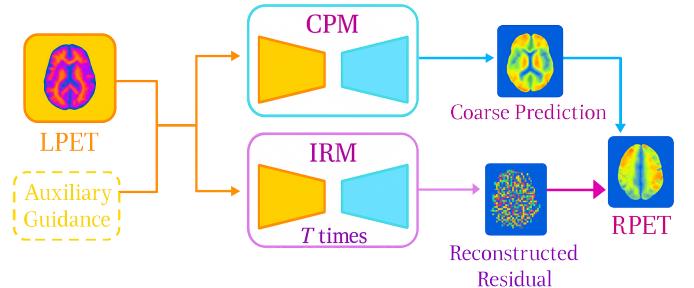
During sampling, starting from Gaussian noise $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, the model iteratively refines the sample through multiple denoising steps. While this produces high-quality results, the iterative nature makes inference computationally expensive compared to single-pass methods like CNNs or GANs.

Diffusion models provide several advantages for PET reconstruction under incomplete ring conditions [18]. They effectively avoid the mode collapse problem common in GANs, ensuring robust coverage of possible solutions [19]. They can also integrate learned priors for PET images, which is crucial when large segments of sinogram data are missing [20]. Their iterative refinement approach aligns well with traditional iterative reconstruction methods in PET imaging, such as MLEM.

However, directly applying DPMs for reconstruction introduces challenges related to computation speed and partial data handling [9]. To address these limitations, we propose 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 [21].

C. Coarse-to-Fine Reconstruction Framework

To effectively address reconstruction challenges in incomplete ring PET imaging, we propose a two-stage approach that balances reconstruction quality with computational efficiency.



quality PET image \mathbf{Y}_B from incomplete data and auxiliary guidance information \mathbf{X}_{aux} . The CPM, implemented as a deterministic U-net, produces a coarse reconstruction \mathbf{x}_{cp} through a single forward pass:

$$\mathbf{x}_{\text{cp}} = P_{\theta}(\mathbf{c}) \quad (10)$$

This module effectively captures global image structures while maintaining computational efficiency [22]. The CPM is optimized using an L1 loss:

$$\mathcal{L}_{\text{CPM}} = \|\mathbf{x}_{\text{cp}} - \mathbf{Y}_A\|_1 \quad (11)$$

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

2. Iterative Refinement Module (IRM)

The second stage employs an Iterative Refinement Module that focuses specifically on recovering residual details. Rather than modeling the entire image, the IRM models only the residual $\mathbf{r}_0 = \mathbf{Y}_A - \mathbf{x}_{\text{cp}}$. This residual-focused approach offers significant advantages:

1. The IRM can utilize a smaller network capacity as it only needs to capture fine details 2. The distribution gap between coarse prediction and ground truth is substantially smaller than between random noise and ground truth 3. 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:

$$\widehat{\mathbf{Y}}_A = \mathbf{x}_{\text{cp}} + \widehat{\mathbf{r}}_0 \quad (12)$$

3. Auxiliary Guidance Strategies

To further improve reconstruction quality, particularly in regions with severe data loss, we incorporate two complementary guidance signals:

Neighboring Axial Slices (NAS). Since PET volumes are inherently three-dimensional, we leverage adjacent slices to provide spatial context [23]:

$$\mathbf{X}_{\text{NAS}} = \{\mathbf{Y}_B^{(z-2)}, \mathbf{Y}_B^{(z-1)}, \mathbf{Y}_B^{(z+1)}, \mathbf{Y}_B^{(z+2)}\} \quad (13)$$

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:

$$\mathbf{X}_{\text{spec}} = \mathcal{F}(\mathbf{Y}_B) \quad (14)$$

Spectral guidance provides global frequency priors that help suppress the high-frequency streak artifacts commonly associated with incomplete angular coverage in PET reconstruction [24].

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

D. Evaluation Metrics

To comprehensively evaluate the performance of incomplete ring PET reconstruction methods, this study employs multiple quantitative and qualitative 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) [25] is a fundamental metric for evaluating reconstructed image quality, calculated based on Mean Squared Error (MSE) and expressed on a logarithmic scale. The definition of PSNR is as follows:

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}_I^2}{\text{MSE}} \right) \quad (15)$$

where 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:

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2 \quad (16)$$

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° - 60°), 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 [26]:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (17)$$

where μ_x and μ_y are the averages of images x and y respectively; σ_x^2 and σ_y^2 are their variances; σ_{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) provides a normalized measure of image error relative to the energy of the original image [27]:

$$\text{NMSE} = \frac{\sum_{i,j,k} (X_{i,j,k} - \hat{X}_{i,j,k})^2}{\sum_{i,j,k} X_{i,j,k}^2} \quad (18)$$

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 10 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 11 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 12 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 (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 13 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, 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 innovation, 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;

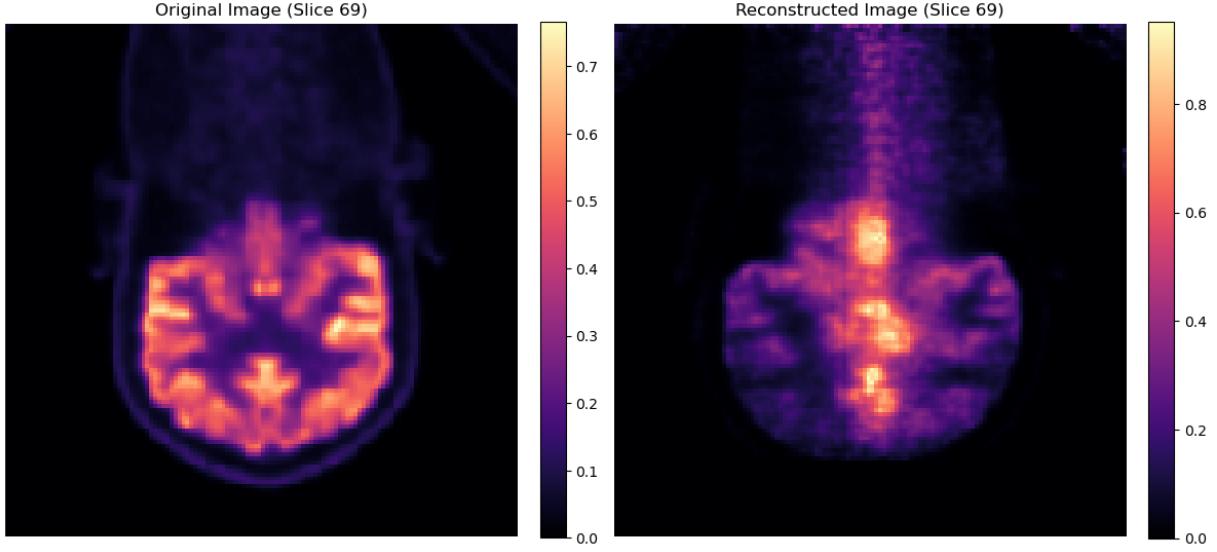


FIG. 10. Comparison of reconstruction results from incomplete rings with the original image, showing that the two are vastly different, demonstrating that direct reconstruction from incomplete rings is not feasible.

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 so-

lution 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).

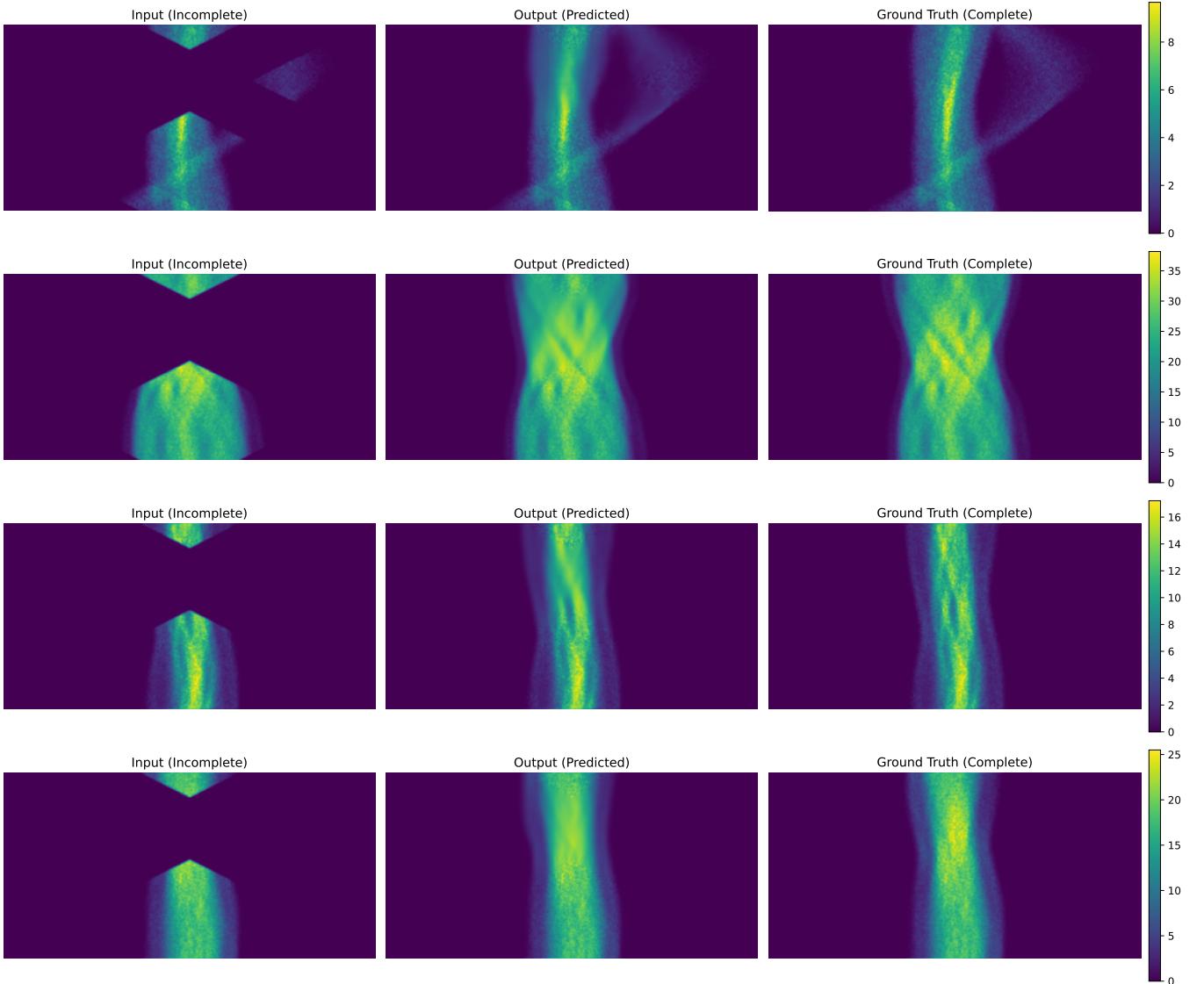


FIG. 11. Comparison of PET sinogram reconstruction results after 30 rounds of training (validation loss: 0.157624). Each row displays three images: the left shows the incomplete input sinogram with missing angular data, the middle shows the model-predicted complete sinogram, and the right shows the complete real sinogram. The color bar indicates signal intensity. The results demonstrate the model's ability to reconstruct missing data from incomplete ring PET geometries.

- [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] 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).
- [11] 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].
- [12] A. J. Reader and B. Pan, Artificial intelligence for PET image reconstruction, *British Journal of Radiology* **96**, 20230292 (2023).
- [13] 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).
- [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] 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.
- [17] J. Ho, A. Jain, and P. Abbeel, Denoising diffusion probabilistic models (2020), arXiv:2006.11239 [cs.LG].

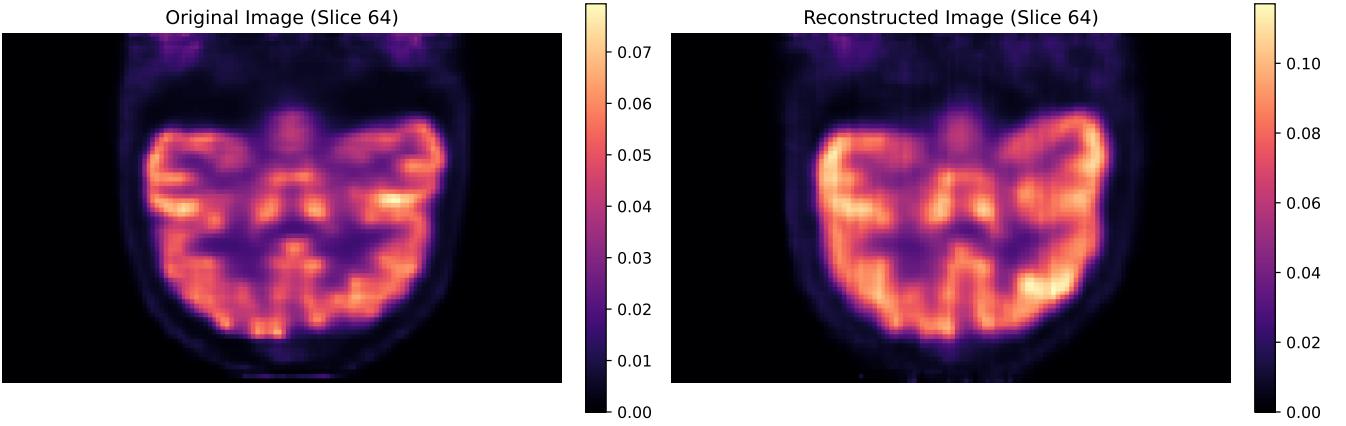


FIG. 12. Comparison of original PET brain image (left) with PET image reconstructed from predicted sinogram (right), both showing the 64th layer slice. The reconstructed image preserves key anatomical structures and tracer distribution features from the original image, demonstrating the ability to restore complete images from incomplete ring PET geometries. The color bar indicates tracer concentration, with the higher maximum value (0.11) in the right reconstructed image compared to the original image (0.07) possibly indicating signal intensity changes during the reconstruction process.

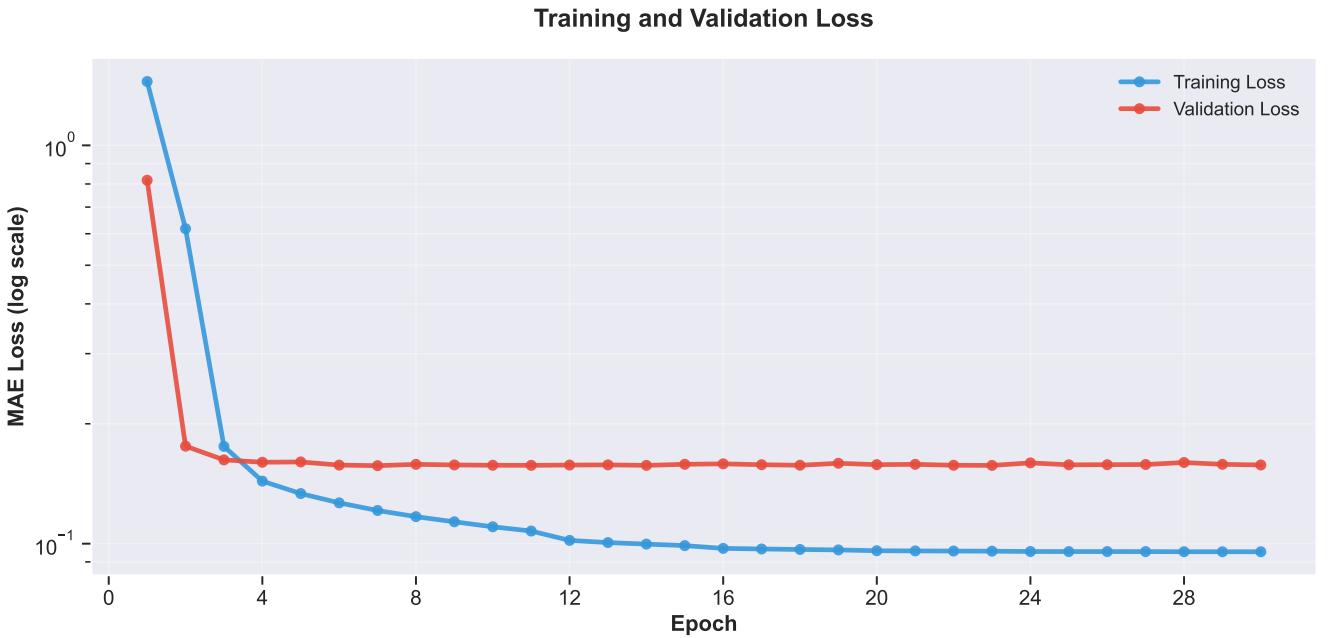


FIG. 13. Change in loss function during the training process of the incomplete ring PET reconstruction model. The figure shows the trends of training loss (blue line) and validation loss (red line) over training rounds, using logarithmic scale Mean Absolute Error (MAE) as the evaluation metric. In the 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.

- [18] G. Webber and A. J. Reader, Diffusion models for medical image reconstruction, *British Journal of Radiology — Artificial Intelligence* **1**, 10.1093/bjrai/ubae013 (2024).
- [19] 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).
- [20] 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).
- [21] 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 In-*

- tervention – MICCAI 2023, LNCS* (Springer, 2023) pp. 239–249.
- [22] 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).
- [23] 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).
- [24] 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.
- [25] 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.
- [26] 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).
- [27] 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).
- [28] F. Neumaier, B. Zlatopolkiy, and B. Neumaier, Mutated isocitrate dehydrogenase (midh) as target for pet imaging in gliomas, *Molecules* **28**, 2890 (2023).
- [29] Ultra-low dose pet imaging challenge 2022, <https://doi.org/10.5281/zenodo.6361846> (2022), accessed: 2025-01-20.
- [30] 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.
- [31] 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.
- [32] 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.
- [33] I. Häggström, C. R. Schmidlein, G. Campanella, and T. J. Fuchs, Deeppet: A deep encoder-decoder network for direct pet image reconstruction, *Medical Image Analysis* **54**, 253 (2019), accessed: 2025-01-20.
- [34] 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.
- [35] 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.
- [36] 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.
- [37] A. Vahdat and J. Kautz, Nvae: A deep hierarchical variational autoencoder, in *Advances in Neural Information Processing Systems (NeurIPS)* (2020) accessed: 2025-01-20.
- [38] 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).
- [39] C.-C. Liu and H.-M. Huang, Partial-ring pet image restoration using a deep learning based method, *Phys. Med. Biol.* **64**, 225014 (2019).
- [40] 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.