



# SCOPE: Slice-COnsistent PET Reconstruction with 2D BBDM

---

2025-Spring DSL Modeling Project Report

**Jeongwoo Lee<sup>†</sup>**  
**Yingjun Shen<sup>‡</sup>**  
**Chaeun Song<sup>‡</sup>**  
**Seungbeom Seo<sup>‡</sup>**

Computer Vision Team

**Data Science Lab, Yonsei**

<sup>†</sup> Project Leader

<sup>‡</sup> Project Members - ordered by cohort and alphabetically

---

## Abstract

Amyloid PET imaging is essential for the diagnosis of Alzheimer's disease, but standard acquisition times of up to 20 minutes can lead to patient discomfort and motion artifacts. In this project, we applied a slice-consistent 2D diffusion model, originally developed for CT-to-MRI translation[3], to reconstruct full-duration PET images from short-time amyloid PET scans. Experimental results demonstrate that it achieves high-quality reconstructions comparable to full-time scans, enabling significant scan time reduction without compromising diagnostic accuracy.

## 1 Introduction

Amyloid PET imaging plays a crucial role in the diagnosis of Alzheimer's disease by visualizing the distribution of amyloid-beta plaques in the brain. However, its widespread clinical utility is hindered by the long acquisition times (typically up to 20 minutes), which can lead to patient discomfort, motion-induced artifacts, and increased costs. Thus, reducing the scan duration while maintaining diagnostic image quality has emerged as a key research challenge.

Recent advances in deep generative models, including GANs[4, 5, 7] and diffusion models[2], have opened new possibilities for reconstructing full-quality PET images from

short-time or low-count acquisitions. While promising, many of these approaches rely on either heavy 3D networks—which demand extensive computational resources—or 2D networks that often produce temporally or anatomically inconsistent slices when stacked into volumes.

In this project, we explore a slice-consistent PET reconstruction diffusion model that incorporates inter-slice spatial context during training. Originally developed for CT-to-MRI translation[3], this architecture combines the efficiency of 2D processing with improved spatial coherence across slices. We adapt this framework for amyloid PET reconstruction and demonstrate that it can effectively reconstruct full-duration scans from only 4-minute acquisitions.

Through quantitative and qualitative evaluations, we show that our method produces reconstructions that are visually and diagnostically comparable to standard full-time scans, paving the way for faster, more comfortable PET imaging without sacrificing accuracy.

## 2 Background

Positron Emission Tomography (PET) is a medical imaging technique that visualizes metabolic activity, blood flow, and glucose consumption in the human body by using radioactive isotopes. It plays a critical role in diagnosing a wide range of diseases, including cancer, Parkinson’s disease, and various neurodegenerative disorders such as dementia.

In a PET scan, a small amount of a radioactive tracer is injected into the patient’s bloodstream. This tracer emits positrons that interact with electrons in the body, producing gamma rays. These gamma rays are detected by the PET scanner to create detailed images of physiological processes. PET provides functional imaging, which complements anatomical imaging methods like MRI or CT, offering deeper insight into disease progression and biological abnormalities.

One specialized application of PET is Amyloid PET, which is used to detect amyloid-beta plaques in the brain - a hallmark of Alzheimer’s disease. Among the commonly used tracers for this purpose is 18F-florbetaben (FBB). Amyloid PET using 18F-FBB allows clinicians to visualize the accumulation of amyloid plaques in brain tissue.

In patients with Alzheimer’s disease, amyloid plaques accumulate abnormally, particularly in the outer cortical regions of the brain. On PET scans, these regions appear bright and white due to the increased uptake of the tracer, providing a visual indication of disease presence and severity. This imaging modality plays a key role in the early and accurate diagnosis of Alzheimer’s, as well as in clinical trials evaluating anti-amyloid therapies.

PET scans, including amyloid PET, typically require prolonged acquisition times (often around 20 minutes) to ensure high image quality and sufficient tracer signal. However, long scan durations present several challenges, such as patient discomfort, motion artifacts, radiation exposure, etc.

## 3 Related Works

Deep generative models such as GANs[4, 5, 7, 8] and diffusion models[2, 3, 6, 9] have increasingly been applied to medical imaging tasks, particularly in scenarios where image quality must be preserved despite constraints like reduced scan time or low signal

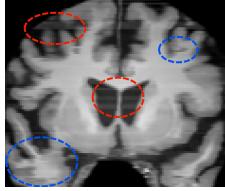


Figure 1: Example of slice inconsistency when using 2D generative models.

counts. In the context of PET image reconstruction, these models have shown promise in synthesizing high-quality images from short-duration acquisitions, offering a potential solution to the long scan time problem inherent in clinical PET imaging.

A wide range of these approaches utilize either 2D[6, 9] or 3D network architectures[8]. 3D generative models can capture inter-slice spatial continuity, often leading to more anatomically consistent reconstructions. However, they tend to be computationally intensive and require substantial resources. In contrast, 2D models are more computationally efficient and easier to train, but they frequently suffer from slice-wise inconsistency, where reconstructed slices may appear misaligned or incoherent when stacked into a volume shown as in Fig. 1.

To address this trade-off, a recent work in CT-to-MRI translation proposed a slice-consistent 2D diffusion model[3] that integrates inter-slice information during training. This approach successfully preserves spatial coherence while maintaining the advantages of 2D architecture. Motivated by this idea, we adopted a similar slice-consistent 2D diffusion framework and extended its application to amyloid PET image reconstruction. Our goal is to synthesize full-duration PET images from 4-minute scans, maintaining both image fidelity and anatomical consistency across slices.

## 4 Dataset & Preprocessing

The dataset utilized for this project was obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI)[1]. It comprised 120 FBB PET scans, each originally acquired as five consecutive 4-minute segments. After filtering out images unsuitable for training due to quality or alignment issues, a total of 54 sets remained, divided into training, validation, and testing subsets with ratios of 38:11:5, respectively. The preprocessing pipeline involved:

1. **Co-registration:** Aligning 4-minute and corresponding full-duration (20-minute) scans to ensure consistent spatial orientation using SPM12.
2. **Voxel Spacing Adjustment:** Standardizing voxel sizes to 1.5mm by merging or splitting voxels to maintain uniform spatial resolution across images.
3. **Normalization:** Adjusting voxel intensities by dividing each voxel value by the average intensity of a reference region (the cerebellum) to ensure comparability across scans.
4. **Smoothing:** Applying a Gaussian smoothing filter (6mm FWHM) to the 4-minute images to reduce noise and minor artifacts.

An example of the preprocessed results is shown in Fig. 2.

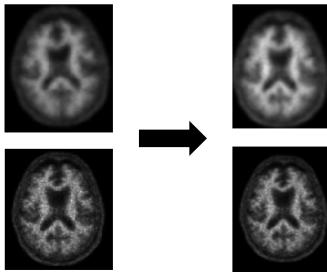


Figure 2: Example of preprocessed results on 4-min and 20-min PET images. Each row shows raw (left) and preprocessed (right) images, with the 4-min scan on top and 20-min on the bottom.

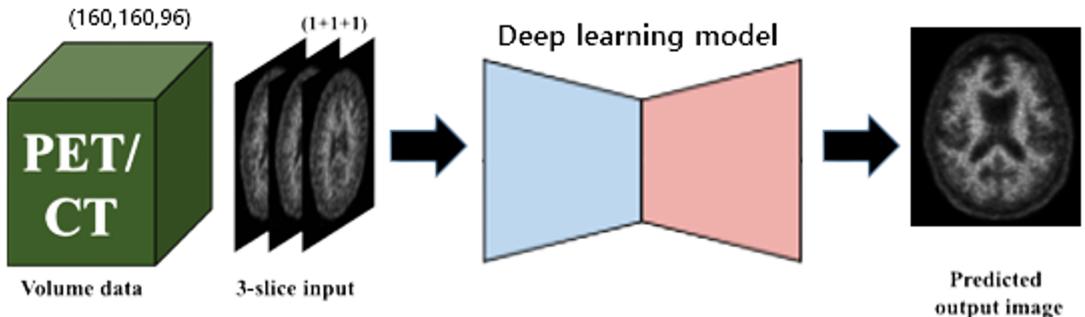


Figure 3: Overall pipeline of the project. A volume of PET data is converted into 3-slice inputs, which are fed into a deep generative model to generate a corresponding output images.

## 5 Pipeline & Methods

In this project, we explored two pipelines for PET image reconstruction from short-duration acquisitions: one based on generative adversarial networks (GANs) and the other on a Brownian Bridge Diffusion Model (BBDM), as illustrated in Fig. 3.

### 5.1 GAN-Based

We experimented with GAN-based models, including three representative architectures: CycleGAN[4], CUT[5], and DCLGAN[7]. These models were applied in a slice-wise manner, where each input consisted of three adjacent PET slices stacked to form a 3-channel image. All GAN-based models were trained for 100 epochs with a batch size of 1 and learning rate of 0.0002.

### 5.2 BBDM-Based

The second pipeline was based on a Brownian Bridge Diffusion Model[2], which operates using a forward and reverse diffusion process with stochastic bridging in the latent space. Similar to the GAN setup, we adopted a slice-wise input format, using three adjacent slices per input.

To enhance slice-consistency, we incorporated two auxiliary techniques used in CT2MRI[3]:

1. **Style Key Conditioning (SKC):** This module guides the denoising process

using the mean intensity histogram of the target PET volume, encouraging stylistic uniformity.

2. **Inter-Slice Trajectory Alignment (ISTA):** This alignment technique encourages trajectory smoothness across adjacent slices during the reverse diffusion process, improving anatomical coherence.

The BBDMs (including baselines) were trained for 50 epochs using default diffusion parameters with a batch size of 1 and learning rate of 0.0001.

## 6 Results

Figure 4 presents a visual comparison of reconstructed PET images generated by various models, including both GAN-based and diffusion-based approaches. The ground truth and input are shown for reference. Some models such as Fast-DDPM[6] and CycleGAN[4] failed to produce meaningful outputs. Among the compared methods, BBDM-based models[2, 3] and DCLGAN[7] visually provide the closest reconstruction to the ground truth.

The performance of both approaches was evaluated using several Image Quality Metrics (IQMs), including NRMSE, PSNR, and SSIM. Table 1 summarizes the reconstruction performance of each model in terms of NRMSE, PSNR, and SSIM. We excluded Fast-DDPM and CycleGAN, since they failed to produce meaningful outputs. Among the evaluated models, BBDM with both SKC and ISTA achieved the best overall performance, recording the lowest NRMSE and the highest PSNR and SSIM scores.

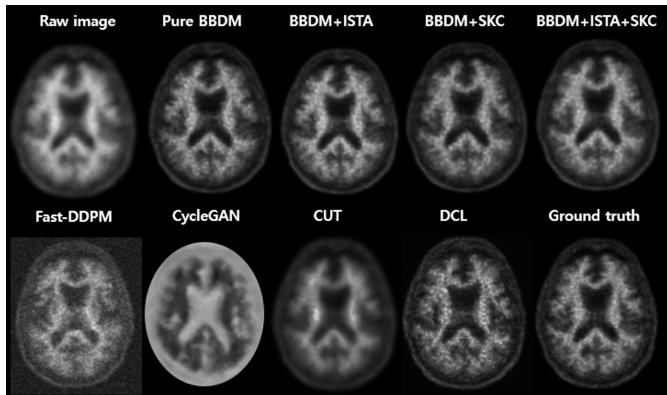


Figure 4: Visual comparison of various generative models.

Table 1: Quantitative comparison of reconstruction quality.

Model	NRMSE $\downarrow$	PSNR $\uparrow$	SSIM $\uparrow$
CUT	0.0629	24.040	0.792
DCLGAN	0.0456	26.756	0.832
BBDM	0.0428	27.479	0.881
BBDM w. SKC	0.0424	27.509	0.895
BBDM w. ISTA	0.0423	27.608	0.894
BBDM w. SKC, ISTA	<b>0.0419</b>	<b>27.673</b>	<b>0.902</b>

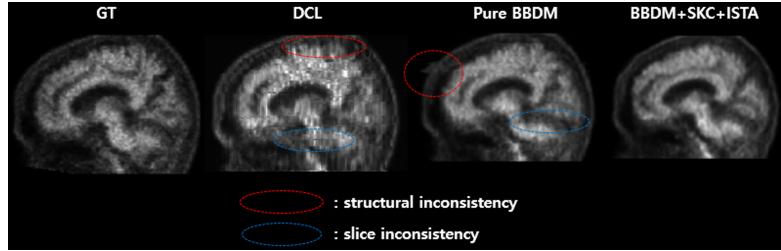


Figure 5: Sagittal-view comparison of various generative models.

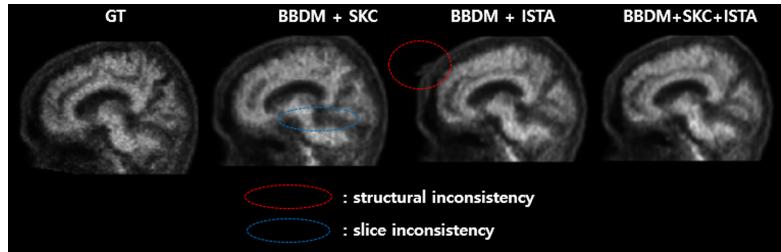


Figure 6: Sagittal-view comparison of different BBDM variants.

In addition to quantitative evaluation, we performed a qualitative analysis to assess the anatomical consistency of reconstructed PET images in the sagittal view. As illustrated in Figure 5 and Figure 6, SKC and ISTA methods contributed significantly to anatomical consistency. Notably, their combination led to the most structurally coherent and slice-consistent reconstructions, closely resembling the ground truth.

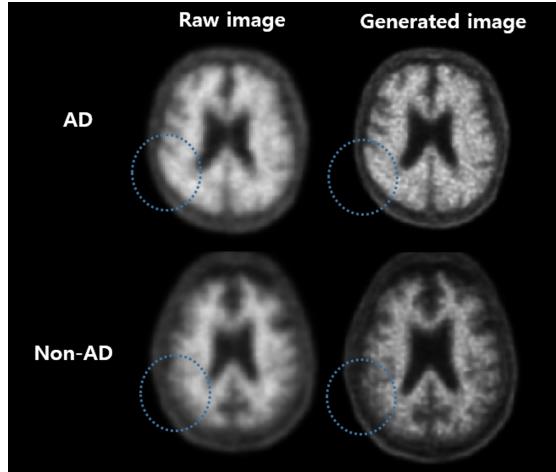


Figure 7: Comparison of raw and generated PET images for AD and non-AD subjects.

Beyond image quality evaluation, we investigated the potential clinical utility of our generated images in the diagnosis of Alzheimer’s disease. In the clinical interpretation of amyloid PET scans, Alzheimer’s disease is often diagnosed based on the relative intensity distribution between the inner white matter and the outer cortical regions. This distinction is difficult to observe in low-quality raw images (4-minute scans). However, with the generated images, the contrast becomes sufficiently clear to distinguish between AD and non-AD cases, as illustrated in 7.

## 7 Conclusions & Limitations

In conclusion, the proposed BBDM-based pipeline effectively reconstructs full-duration amyloid PET scans from significantly shorter acquisitions (4-minute scans). The integration of SKC and ISTA techniques successfully mitigated slice inconsistency issues typical of 2D approaches.

However, several limitations remain:

1. **Limited Dataset Size:** The relatively small number of available scans (54 usable samples) may restrict generalizability.
2. **Alzheimer’s Data Scarcity:** The dataset contained relatively few confirmed Alzheimer’s cases, which could limit the effectiveness of the model in diverse clinical settings.

**Future Directions.** To address these limitations, future work may focus on expanding the dataset to include a more diverse and balanced cohort, particularly with a greater number of AD-positive cases. Beyond BBDM, applying the proposed SKC and ISTA techniques to other 2D diffusion-based generative models could be explored to assess the generalizability and effectiveness of these methods across architectures.

## References

- [1] Alzheimer’s Disease Neuroimaging Initiative (ADNI).  
<https://adni.loni.usc.edu/>
- [2] Li, Bo and Xue, Kaitao and Liu, Bin and Lai, Yu-Kun. Image-to-image Translation with Brownian Bridge Diffusion Models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023.
- [3] Kyobin Choo and Youngjun Jun and Mijin Yun and Seong Jae Hwang. Slice-Consistent 3D Volumetric Brain CT-to-MRI Translation with 2D Brownian Bridge Diffusion Model. *MICCAI*, 2024.
- [4] Jun-Yan Zhu and Taesung Park and Phillip Isola and Alexei A. Efros. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. *ICCV*, 2017.
- [5] Taesung Park and Alexei A. Efros and Richard Zhang and Jun-Yan Zhu. Contrastive Learning for Unpaired Image-to-Image Translation. *ECCV*, 2020.
- [6] Hongxu Jiang and Muhammad Imran and Linhai Ma and Teng Zhang and Yuyin Zhou and Muxuan Liang and Kuang Gong and Wei Shao. Fast-DDPM: Fast Denoising Diffusion Probabilistic Models for Medical Image-to-Image Generation. *JBHI*, 2025.
- [7] Junlin Han and Mehrdad Shoeiby and Lars Petersson and Mohammad Ali Armin. Dual Contrastive Learning for Unsupervised Image-to-Image Translation. *NTIRE, CVPRW*, 2021.
- [8] Huidong Xie and Weijie Gan and Bo Zhou and Ming-Kai Chen and Michal Kulon and Annemarie Boustani and Benjamin A. Spencer and Reimund Bayerlein and Wei Ji and Xiongchao Chen and Qiong Liu and Xueqi Guo and Menghua Xia and Yinchi

Zhou and Hui Liu and Liang Guo and Hongyu An and Ulugbek S. Kamilov and Hanzhong Wang and Biao Li and Axel Rominger and Kuangyu Shi and Ge Wang and Ramsey D. Badawi and Chi Liu. 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.

- [9] Zeyu Han and Yuhan Wang and Luping Zhou and Peng Wang and Binyu Yan and Jiliu Zhou and Yan Wang and Dinggang Shen. Contrastive Diffusion Model with Auxiliary Guidance for Coarse-to-Fine PET Reconstruction. *MICCAI*, 2023.