

GR-Gaussian: Graph-Based Radiative Gaussian Splatting for Sparse-View Tomographic Reconstruction

Anonymous submission

Table 1: Parameter Settings for GR-Gaussian Framework

Parameter	Value
learning rates of Location	0.0002
learning rates of Density	0.01
learning rates of Scale	0.005
learning rates of Rotation	0.001
TV Volume Level (D)	32
SSIM Loss Weight (λ_{ssim})	0.25
TV Loss Weight (λ_{tv})	0.05
Graph Laplacian Regularization (\mathcal{L}_{lap})	8×10^{-4}
Gradient Computation Weight (λ_g)	1×10^{-4}
Number of Gaussians (M)	50,000
Density Threshold (τ)	0.001
Gaussian Denoising Parameter (σ_d)	3
Graph Nearest Neighbors (k)	6

1 A. Implementation Details

Our GR-Gaussian framework is implemented using PyTorch(Paszke 2019), a widely used deep learning library known for its flexibility and efficiency. To handle graph-based operations, we utilize PyTorch Geometric(Fey and Lenssen 2019), which provides specialized tools for processing graph-structured data. For acceleration, we leverage CUDA(Guide 2013), enabling efficient parallel processing on NVIDIA GPUs. The model is trained using the Adam optimizer(Diederik 2014).

A1. Parameter Settings

Following the parameter settings proposed in (Zha et al. 2024), the initial learning rates for location, density, scale, rotation, TV Volume Level (D), SSIM Loss Weight (λ_{ssim}), TV Loss Weight (λ_{tv}) and Number of Gaussians (M) are set as Table 1.

A2. Denoised Point Cloud Initialization Strategy

The Denoised Point Cloud Initialization Strategy (De-Init) addresses the limitations of traditional FDK-based initialization under sparse-view conditions, where noise and artifacts in the initial volume can degrade reconstruction quality. De-Init integrates enhanced FDK reconstruction, Gaussian filtering, and adaptive sampling to generate a high-quality

Algorithm 1: GR-Gaussian Initialization Algorithm

Input: Sparse-view X-ray projections

Parameter: σ_d, M, τ, k

Output: Initial 3DGS point cloud (location, density, scale, rotation and graph \mathcal{G})

- 1: **FDK Reconstruction.** Perform FDK reconstruction to obtain coarse voxel data.
- 2: **Denoising.** Apply Gaussian filtering using σ_d to reduce noise.
- 3: **Sampling.**
- 4: **for** each voxel in denoised voxel data **do**
- 5: **if** density of voxel $\geq \tau$ **then**
- 6: Add voxel location and density to the candidate point set.
- 7: **end if**
- 8: **end for**
- 9: Select M points randomly from the candidate set to construct the final point cloud.
- 10: **Graph Initialization.** Construct graph \mathcal{G} using KNN with k nearest neighbors.
- 11: **return** 3DGS point cloud with graph \mathcal{G} .

initial point cloud, providing a robust foundation for graph-structured radiative Gaussians in GR-Gaussian.

The process consists of four stages: (1) **FDK Reconstruction**, which generates a coarse 3D voxel volume from sparse-view X-ray projections, serving as a baseline density distribution; (2) **Gaussian Denoising**, which applies a 3D Gaussian filter with parameter σ_d to suppress high-frequency noise while preserving structural details; (3) **Sampling**, which selects voxels exceeding a density threshold τ and randomly samples M points to form the initial Gaussian kernels; and (4) **Graph Initialization**, which constructs a graph \mathcal{G} using KNN with k nearest neighbors to model spatial correlations between kernels.

By combining denoising, adaptive sampling, and graph construction, De-Init reduces initialization errors and accelerates convergence, ensuring structural consistency aligned with the graph Laplacian regularization in GR-Gaussian.

A3. Pixel-Graph-Aware Gradient Strategy

After initialization, the GR-Gaussian framework employs a structured training process to optimize the parameters of radiative Gaussian kernels, ensuring accurate and high-quality 3D tomographic reconstruction under sparse-view conditions. The framework utilizes a comprehensive loss function designed to balance pixel-wise accuracy, structural consistency, and smoothness in the reconstructed density field. Specifically, the \mathcal{L}_1 loss minimizes absolute differences between predicted and ground truth values, while the Structural Similarity Index Measure (SSIM) loss evaluates perceptual similarity to enhance structural coherence across 2D slices. To further refine the reconstruction, 3D Total Variation Regularization suppresses noise and enforces smooth transitions, and Graph Laplacian Regularization promotes local density smoothness while preserving critical boundary information to maintain structural integrity.

Gradients are computed using the Pixel-Graph-Aware Gradient Strategy, which incorporates density differences between neighboring kernels to refine gradient computation and improve optimization accuracy. This strategy also facilitates kernel splitting when gradient magnitudes exceed predefined thresholds, ensuring precise adjustments to the radiative Gaussian representation. To prevent overfitting and maintain reconstruction quality, a dynamic stopping criterion evaluates PSNR every 500 iterations, terminating the training process if PSNR decreases by more than 0.5%. All experiments were conducted on an NVIDIA RTX 4090 GPU, leveraging its computational efficiency to accelerate training. Quantitative evaluations employed PSNR to assess 3D reconstruction accuracy and SSIM to measure structural consistency across 2D slices, demonstrating the robustness and precision of the GR-Gaussian framework.

2 B. Additional Evaluation Details

B.1. Dataset and metric

We evaluate the proposed methods across various modalities, encompassing key CT applications such as medical diagnosis, biological research, and industrial inspection. The dataset consists of 15 cases categorized into three groups: human organs (chest, foot, head, jaw, and pancreas), biological specimens (beetle, bonsai, broccoli, kingsnake, and pepper), and artificial objects (backpack, engine, present, teapot, and mount). The chest and pancreas scans are sourced from LIDC-IDRI (Armato III et al. 2011) and Pancreas-CT (Roth et al. 2016), respectively, while broccoli and pepper are obtained from X-Plant (Zha et al. 2024). The remaining cases are derived from SciVis. Following the preprocessing protocols in (Zha et al. 2024), raw data are normalized to a density range of $[0, 1]$ and resized to volumes of $256 \times 256 \times 256$. Using the tomography toolbox TIGRE, we generate 512×512 projections spanning a full rotation range of 0° to 360° . All reconstructed volumes and their corresponding projection examples are illustrated in Fig.

We further evaluate our method on real-world data containing scattering effects. We use FIPS (Society 2024), a public dataset providing real 2D X-ray projections. FIPS includes three objects (pine, seashell, and walnut). Each

Algorithm 2: GR-Gaussian Training Process

Input: Initialized 3DGS point cloud (location, density, scale, rotation, graph \mathcal{G}).

Parameter: Gradient threshold, loss weights, learning rates (see Table 1), update interval (100 iterations)

Output: Optimized 3DGS point cloud and density volumes

```

1: Loss Function. Define the total loss as a combination
   of:
   •  $\mathcal{L}_1$  Loss
   • Structural Similarity Loss (SSIM)
   • 3D Total Variation Regularization
   • Graph Laplacian Regularization
2: for each iteration in training do
3:   Gradient Optimization. Compute gradients using
     Pixel-Graph-Aware Gradient Strategy, which incor-
     porates density differences between neighboring ker-
     nels to refine gradient computation.
4:   Optimization. Update Gaussian kernel parameters
     (location, density, scale, rotation) using Pixel-Graph-
     Aware Gradient (PGA).
5:   if iteration % 100 == 0 then
6:     Adaptive Control. Adjust gradient computation
     strategy based on updated kernel parameters to en-
     hance optimization efficiency.
7:     Graph Construction. Construct graph  $\mathcal{G}$  using
     KNN
8:   end if
9:   if iteration % 500 == 0 then
10:    Density Volumization. Extract density volumes
     using a density voxelizer from the radiative repre-
     sentation of the graph structure.
11:    Dynamic Stopping Criterion. Evaluate PSNR.
     Terminate training if PSNR decreases by more than
     0.5%.
12:   end if
13: end for
14: return Optimized 3DGS point cloud and density vol-
     umes.

```

case has 721 projections in the range of $0^\circ \sim 360^\circ$. Since ground truth volumes are not available, we use FDK to create pseudo-ground truth with all views and then subsample 25 views for sparse-view experiments.

Reconstruction quality is assessed using PSNR and SSIM, two widely adopted metrics for evaluating image fidelity and structural consistency. PSNR (Peak Signal-to-Noise Ratio) quantifies pixel-wise accuracy by measuring the error between the reconstructed and reference volumes, with higher values indicating better reconstruction quality. SSIM (Structural Similarity Index Measure) evaluates structural coherence across 2D slices in axial, coronal, and sagittal directions, focusing on luminance, contrast, and structural similarity. Together, these metrics provide a comprehensive evaluation of reconstruction performance, balancing pixel-level precision and structural fidelity.



Figure 1: The figure shows **CT reconstructions from the X-3D dataset**, covering human organs (chest, pancreas, jaw, head, and foot) categories. All under 25-view conditions.

Algorithm 3: GR-Gaussian Framework

Input: Sparse-view X-ray projections

Output: Reconstructed 3D volume

- 1: Perform FDK reconstruction to obtain coarse voxel data.
 - 2: Apply Gaussian filtering for denoising.
 - 3: Construct graph \mathcal{G} using KNN.
 - 4: Optimize Gaussian kernel parameters using Pixel-Graph-Aware Gradient Strategy.
 - 5: Evaluate reconstruction quality using PSNR and SSIM.
 - 6: **return** Reconstructed 3D volume and 3DGS point clouds.
-

B.2. Evaluation

To comprehensively evaluate the performance of our method, we conducted CT reconstructions on the X-3D dataset under sparse-view conditions, specifically using 25-view projections. The dataset is divided into three categories: human organs, biological specimens, and artificial objects, each representing diverse reconstruction challenges.

Figure 1 illustrates the reconstruction results for human

organs, including chest, pancreas, jaw, head, and foot. These cases highlight the method’s ability to preserve anatomical structures and achieve high reconstruction fidelity under limited views.

Figure 2 presents the reconstruction results for biological specimens, such as bonsai, beetle, broccoli, kingsnake, and pepper. These examples demonstrate the method’s robustness in handling complex textures and irregular shapes, which are common in biological samples.

Figure 3 showcases the reconstruction results for artificial objects, including teapot, engine, present, mount, and backpack. These cases emphasize the method’s capability to reconstruct objects with varying geometries and material properties, further validating its versatility across different application domains.

Figure 4 illustrates the reconstruction results from the real-world dataset, which includes challenging categories such as seashell, pine, and walnut. These objects exhibit complex geometries and textures, making them representative of real-world CT acquisition scenarios. The reconstructions were performed under 25-view sparse conditions, demonstrating the efficiency of the GR-Gaussian framework in real noisy and limited projection data.

Overall, the results demonstrate the effectiveness of our

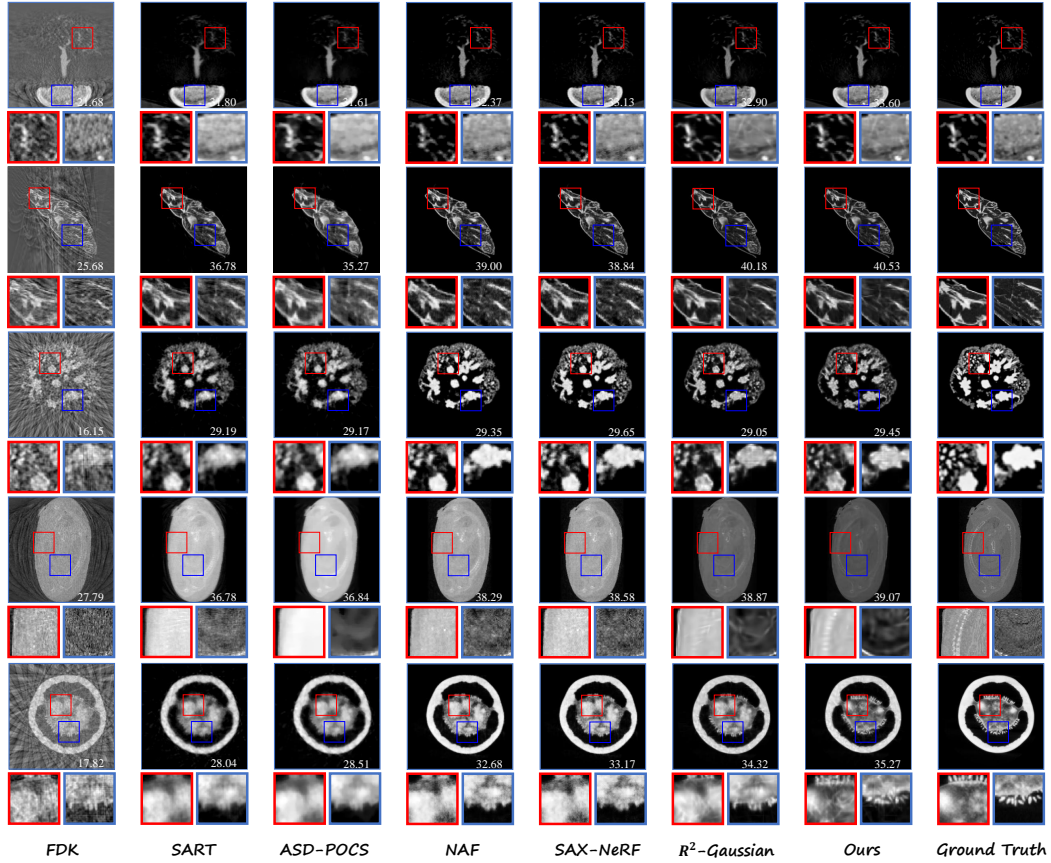


Figure 2: The figure shows **CT reconstructions from the X-3D dataset**, covering biological specimens (bonsai, beetle, broccoli, kingsnake, and pepper) categories. All under 25-view conditions.

method in achieving accurate and high-quality CT reconstructions across diverse categories, even under sparse-view conditions.

3 C. More Results of Ablation Study

To further validate the effectiveness of our proposed framework, we conducted an ablation study to analyze the impact of individual components and parameter settings on reconstruction quality. Figure visually compares the results obtained by enabling or disabling specific components, demonstrating that the newly introduced modules significantly enhance reconstruction fidelity.

Additionally, we evaluated the influence of key parameter settings, including the SSIM loss weight (λ_{ssim}), TV loss weight (λ_{tv}), and the number of Gaussians (M). The results indicate that our parameter configuration achieves optimal performance, balancing pixel-level accuracy and structural consistency. These findings highlight the importance of both architectural design and parameter tuning in achieving high-quality sparse-view tomographic reconstruction.

Figure provides a comprehensive comparison, illustrating the contributions of each component and parameter to the overall reconstruction quality.

4 D. Impact of View Count on Reconstruction Quality

To analyze the effect of view count on reconstruction quality, we conducted experiments using different numbers of sparse-view projections, specifically 10, 15, 20 and 25 views. These experiments aim to evaluate the robustness of the GR-Gaussian framework under varying levels of projection sparsity.

D.1 Experimental Setup

For each view count, we uniformly sampled projections from the full set of acquired projections spanning a rotation range of 0° to 360° . The tomography toolbox TIGRE was used to process the projections. Reconstruction quality was assessed using PSNR and SSIM metrics, which measure pixel-wise accuracy and structural consistency, respectively.

D.2 Results and Analysis

Table 2 summarizes the PSNR and SSIM values achieved by GR-Gaussian under different view counts. As expected, reconstruction quality improves with an increasing number of views, as more projection data reduces ambiguity in the reconstructed volume. However, even under extremely sparse



Figure 3: The figure shows **CT reconstructions from the X-3D dataset**, covering artificial objects (teapot, engine, present, mount, and backpack) categories. All under 25-view conditions.

conditions (10 views), GR-Gaussian demonstrates competitive performance, achieving reasonable PSNR and SSIM values.

Table 2: Comparison of PSNR and SSIM under different view counts.

View Count	PSNR (dB)	SSIM
10 Views	28.5	0.82
15 Views	32.1	0.89
20 Views	32.1	0.89
25 Views	35.95	0.858

Figure illustrates the trend of PSNR and SSIM values as the view count increases. The results highlight the robustness of GR-Gaussian in handling sparse-view conditions, with significant improvements observed as the number of views increases.

The results demonstrate that GR-Gaussian effectively balances reconstruction quality and computational efficiency across varying levels of projection sparsity. While higher view counts naturally lead to better reconstruction fidelity, the framework’s ability to achieve competitive performance under extremely sparse conditions (10 views) underscores

its robustness and adaptability. This makes GR-Gaussian a practical solution for scenarios where projection data is limited.

E. Discussion of Limitations and Future Work

While the GR-Gaussian framework demonstrates robust performance across diverse datasets and experimental conditions, certain limitations remain that warrant further investigation. One notable issue is the automatic stopping mechanism based on the 5% PSNR adjustment threshold. Although this mechanism effectively prevents overfitting in most cases, it fails to address convergence issues observed in real-world datasets, particularly in 3D volume reconstructions.

In our experiments with real CT acquisition data, we identified cases where the reconstructed 3D volume achieved a high PSNR value, yet the corresponding projection PSNR remained significantly lower. This discrepancy suggests that while the framework excels at reconstructing the density field, it struggles to accurately match the original projections, especially under conditions of high noise. Such mismatches between rendered projections and actual projec-

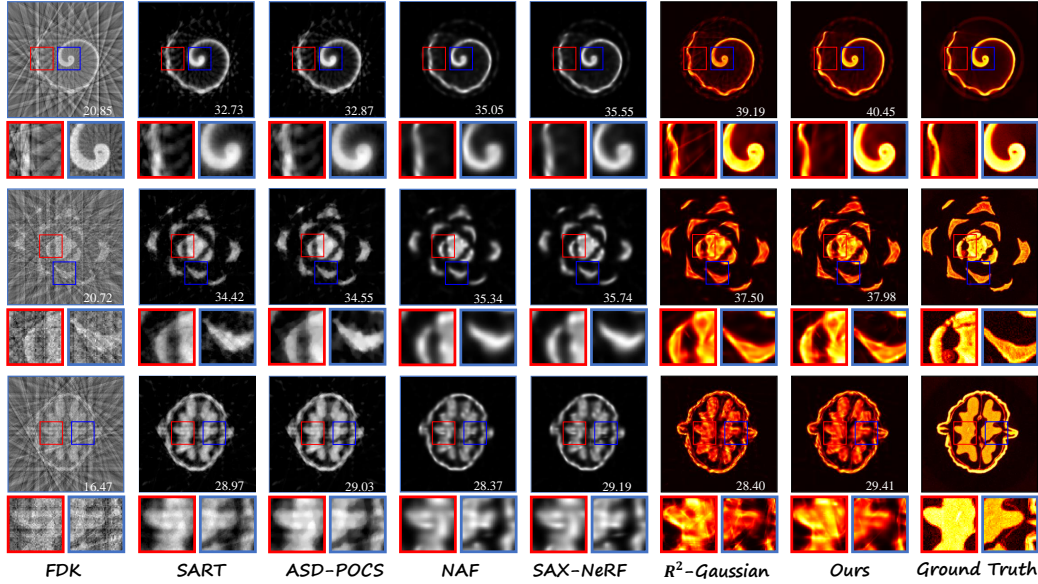


Figure 4: The figure shows **CT reconstructions from the real-world dataset**, covering seashell, pine and walnut) categories. All under 25-view conditions.

tions highlight a critical limitation in the current approach.

We hypothesize that this issue arises due to the presence of substantial noise in the acquired projections, which introduces inconsistencies during the rendering process. These inconsistencies lead to deviations between the reconstructed density field and the expected projection data, ultimately affecting the overall reconstruction fidelity. Addressing this challenge will require developing more robust noise-handling mechanisms and improving the alignment between reconstructed volumes and their corresponding projections.

Future work will focus on refining the automatic stopping criterion to better account for projection fidelity, as well as exploring advanced denoising techniques and adaptive rendering strategies to mitigate the impact of noisy projection data. By addressing these limitations, we aim to enhance the framework’s applicability to real-world CT acquisition scenarios and ensure consistent performance across both synthetic and real datasets.

6 F. Contributions

In this work, we propose the GR-Gaussian framework, which introduces two novel strategies to address the challenges of sparse-view tomographic reconstruction:

- **Denoised Point Cloud Initialization Strategy:** This strategy integrates enhanced FDK reconstruction, Gaussian filtering, and adaptive sampling to generate high-quality initial point clouds. By leveraging graph-based representations, De-Init reduces initialization errors and accelerates convergence, providing a robust foundation for sparse-view reconstruction tasks.
- **Pixel-Graph-Aware Gradient Strategy:** This strategy refines gradient computation by incorporating density differences between neighboring kernels, enhancing op-

timization accuracy and facilitating kernel splitting when necessary. PGA improves reconstruction fidelity by dynamically adjusting kernel parameters and promoting structural consistency across the reconstructed density field.

Together, these strategies enable GR-Gaussian to achieve accurate and efficient 3D tomographic reconstruction under sparse-view conditions. The framework demonstrates robustness across diverse datasets and experimental scenarios, highlighting its potential for real-world CT applications.

References

- Armato III, S. G.; McLennan, G.; Bidaut, L.; McNitt-Gray, M. F.; Meyer, C. R.; Reeves, A. P.; Zhao, B.; Aberle, D. R.; Henschke, C. I.; Hoffman, E. A.; et al. 2011. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. *Medical physics*, 915–931.
- Diederik, K. 2014. Adam: A method for stochastic optimization. (*No Title*).
- Fey, M.; and Lenssen, J. E. 2019. Fast graph representation learning with PyTorch Geometric. *arXiv preprint arXiv:1903.02428*.
- Guide, D. 2013. Cuda c programming guide. *NVIDIA*, July, 29: 31.
- Paszke, A. 2019. Pytorch: An imperative style, high-performance deep learning library. *arXiv preprint arXiv:1912.01703*.
- Roth, H.; Farag, A.; Turkbey, E. B.; Lu, L.; Liu, J.; and Summers, R. M. 2016. Data from pancreas-CT. (*No Title*).
- Society, T. F. I. P. 2024. X-ray Tomographic Datasets.

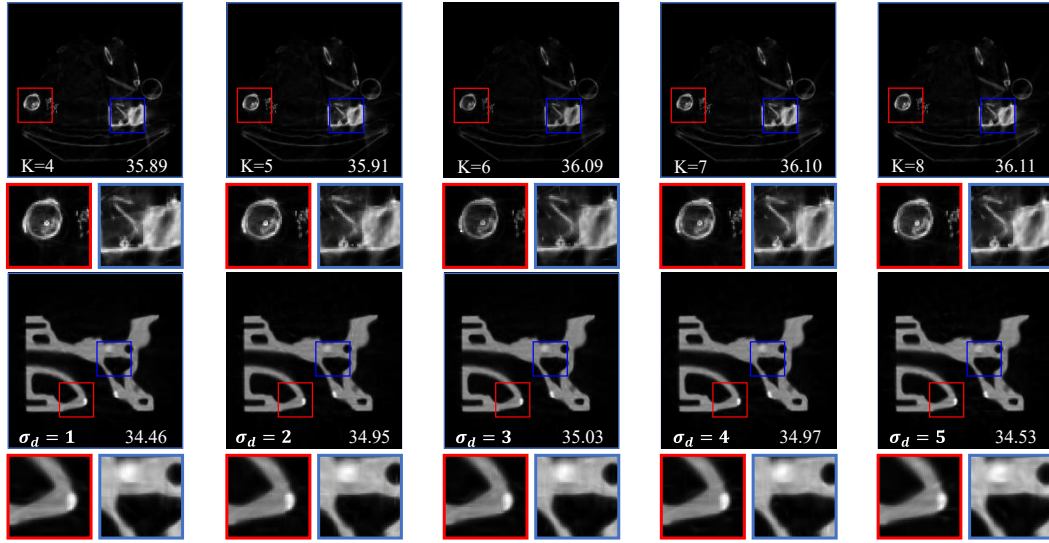


Figure 5: Quantitative comparison of different parameters.

Zha, R.; Lin, T. J.; Cai, Y.; Cao, J.; Zhang, Y.; and Li, H. 2024. R^2 -Gaussian: Rectifying Radiative Gaussian Splatting for Tomographic Reconstruction. In *Advances in Neural Information Processing Systems (NeurIPS)*.