UniSpine-GS: An Efficient Physics-Aware Gaussian Framework for Cross-Modality Multi-view Spine Image Synthesis

Qiuhua Chen[1], Changning Yu[1], Na Huang[2], Chao Sun[1,3](✉), and Bo Du[1,3,4,5]

[1] School of Computer Science, Wuhan University, Wuhan, China
[2] The Department of Ultrasound of Renmin Hospital East Branch of Wuhan University, Wuhan, China
[3] Institute of Artificial Intelligence, Wuhan University, Wuhan, China
[4] National Engineering Research Center for Multimedia Software, Wuhan University, Wuhan, China
[5] Hubei Key Laboratory of Multimedia and Network Communication Engineering, Wuhan University, Wuhan, China
chaosun@whu.edu.cn

Abstract. The diagnosis of spinal diseases is often assisted by 3D imaging techniques in clinical practice. However, precise 3D spinal assessment is limited by the high costs of 3D imaging hardware and the challenges posed by the physical differences between imaging modalities, which hinder the generalizability of models. To address these issues, we propose **UniSpine-GS**, an efficient, physics-aware Gaussian framework designed for novel-view projection rendering in multi-view spine imaging via a 3D-aware representation. Instead of performing explicit 3D reconstruction, our approach learns a geometry-aware Gaussian representation that ensures anatomical consistency across different views. We introduce SPWM, a structure-guided loss reweighting strategy to improve boundary fidelity and local details. We evaluate our method on the CTSpine3D dataset and a newly constructed 3D fetal ultrasound dataset, FeSpine3D. Our results demonstrate that UniSpine-GS significantly outperforms existing methods across all metrics, offering a practical and cost-effective solution for unified multi-view medical imaging. Our code is publicly available at https://github.com/orangeisland66/UniSpine-GS.

Keywords: Multi-view Spine Imaging · Cross-Modality 3D Synthesis · X-Gaussian Representation · Medical Image Synthesis

1 Introduction

Spinal disorders such as scoliosis, vertebral degeneration, and spondylolisthesis require an accurate assessment of spinal anatomy in three dimensions (3D) for reliable diagnosis and treatment planning [13]. In clinical practice, X-ray imaging is widely used due to its low cost, low dose, and broad accessibility, while ultrasound is highly valuable for bedside examinations and obstetric scenarios because it is radiation-free, portable, and real-time. However, accurate 3D spinal assessment remains challenging. High-quality 3D imaging and dedicated acquisition systems are expensive and not always available. Moreover, large differences in imaging mechanisms and visual appearance across modalities make it difficult for models to maintain structural consistency and generalization in multi-modality and multi-view settings [11, 26]. These limitations motivate the need for cost-effective methods that enable cross-view and cross-modality spine image synthesis with reliable structural representations.

Recent advances in neural fields and differentiable rendering have enabled learning 3D representations from limited views for novel-view synthesis. NeRF [16] achieves high fidelity but remains computationally expensive, motivating accelerated neural-field variants with efficient encodings and anti-aliasing designs [2,17]. In contrast, 3D Gaussian Splatting (3DGS) uses explicit Gaussian primitives and parallel rasterization for efficient training and rendering [1, 12]. For projection imaging, radiative Gaussian methods combine radiative forward models with Gaussian representations [4] and have been extended to rectified and continuous-time 4D tomographic reconstruction [21,24], with structure-aware priors improving sparse-view stability [5]. Gaussian representations have also been explored for SLAM, dynamic scenes, inverse rendering [14, 19, 22], SfM-free initialization, compact parameterizations [7, 10], and generative modeling from limited observations [18], while attention- or modality-enhanced neural fields improve robustness under non-ideal settings [9, 15]. Nevertheless, cross-modality multi-view spine synthesis remains underexplored, especially for ultrasound, whose reflection/scattering-driven formation differs fundamentally from X-ray transmission and makes faithful acoustic simulation costly.

Building on these ideas, we propose UniSpine-GS, an efficient and physics-aware Gaussian framework for cross-modality, multi-view 2D spine image synthesis. UniSpine-GS optimizes a geometry-aware Gaussian representation and enforces cross-view consistency through a unified forward operator. For ultrasound, we use a pragmatic proxy by treating the 3D volume as a projectable scalar field and generating multi-view images using the same differentiable digitally reconstructed radiograph (DRR) operator. To improve fine-grained anatomy under sparse views, we introduce a Structure Prior Weight Map (SPWM) that reweights the loss to emphasize vertebral boundaries and texture-informative regions, with warm-up scheduling for stable optimization.

Our contributions are summarized as follows:

- We propose UniSpine-GS, an efficient physics-aware Gaussian framework for cross-modality multi-view 2D spine image synthesis.
- We introduce an SPWM, a parameter-light structure-guided loss reweighting strategy with warm-up scheduling to enhance boundary fidelity and local details under sparse-view supervision.
- We construct the FeSpine3D dataset and validate UniSpine-GS on CTSpine3D and FeSpine3D, achieving consistent improvements in structural consistency and local detail with high efficiency.

2 Methodology

The overall architecture of UniSpine-GS is a differentiable radiative rendering pipeline with an explicit 3D Gaussian representation. Given multi-view observations and acquisition parameters, we initialize and optimize a set of Gaussians and render a projection via a radiative forward operator. Training minimizes a hybrid objective combining an SSIM loss and an SPWM-reweighted Charbonnier loss. Gaussian parameters are updated through backpropagation with a densification–pruning schedule to progressively refine the representation.

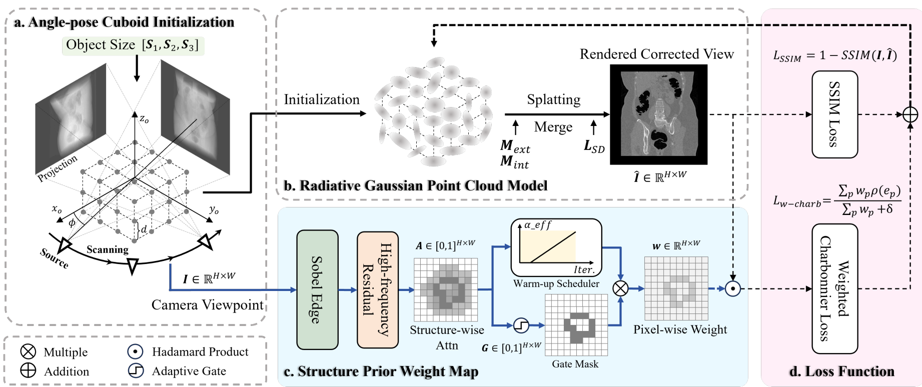

Fig. 1. Overview of UniSpine-GS. We optimize a radiative Gaussian representation for multi-view spine synthesis using a unified forward operator. An SPWM derived from the Camera Viewpoint image reweights the reconstruction loss, together with an SSIM term, to emphasize structurally informative regions.

2.1 Angle-pose Cuboid Initialization (ACUI)

Efficient optimization requires reliable camera-pose and Gaussian-center initialization. Unlike natural-scene 3DGS relying on SfM, projection imaging provides scanner geometry, whereas low contrast and signal superposition hinder feature matching. We adopt ACUI from X-Gaussian [4] for circular cone-beam acquisition, deriving camera parameters and projection matrices from scanner geometry and azimuth ϕ. Gaussian centers are uniformly sampled at interval d inside an origin-centered cuboid enclosing the anatomy, with dimensions $\boldsymbol{S}_1$, $\boldsymbol{S}_2$, and $\boldsymbol{S}_3$. This initialization provides a spatial prior and improves sparse-view stability.

2.2 Radiative Gaussian Point Cloud Model

Inspired by X-Gaussian [4], our model formulates radiative Gaussian splatting for projection imaging. An object is represented as a set of 3D Gaussian primitives:

$$\mathcal{G} = \{(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i, \alpha_i, \mathbf{f}_i)\}_{i=1}^{N_p}, \tag{1}$$

where $\boldsymbol{\mu}_i$ denotes the Gaussian center, $\boldsymbol{\Sigma}_i$ the covariance, α_i the opacity, and $\mathbf{f}_i$ the learnable radiative feature.

Following X-Gaussian, we model a view-independent radiation intensity using a Radiation Intensity Response Function (RIRF):

$$\mathbf{i}_i = \mathrm{Sigmoid}(\boldsymbol{\lambda} \odot \mathbf{f}_i), \tag{2}$$

where $\boldsymbol{\lambda}$ is a fixed weight vector and $\odot$ denotes element-wise multiplication.

Given camera extrinsics $\mathbf{M}_{ext}$ and intrinsics $\mathbf{M}_{int}$, we render the projection via a differentiable radiative forward operator based on transmittance-driven compositing along each ray:

$$\hat{\boldsymbol{I}} = F_{\mathrm{DRR}}(\mathbf{M}_{ext}, \mathbf{M}_{int}, \mathcal{G}), \tag{3}$$

$$\hat{\boldsymbol{I}}(p) = \sum_j T_j(p)\, \sigma_j(p)\, \mathbf{i}_j, \tag{4}$$

$$T_j(p) = \prod_{k<j} \big(1 - \sigma_k(p)\big), \tag{5}$$

where $\sigma_j(p)$ denotes the contribution of the j-th Gaussian at pixel p, and $T_j(p)$ is the accumulated transmittance. This radiative formulation enables efficient, physically consistent forward modeling [4, 24], while UniSpine-GS further introduces a structure-aware strategy to enhance boundary fidelity and consistency.

2.3 Structure Prior Weight Map (SPWM)

To improve fine-grained anatomical reconstruction, we introduce an SPWM that emphasizes informative regions during training.

Structure-wise Attention Map Given the ground-truth (GT) projection used only during training $\boldsymbol{I} \in \mathbb{R}^{H\times W}$, we compute a parameter-light Structure-wise Attention Map $\boldsymbol{A} \in [0, 1]^{H\times W}$. We extract Sobel gradients and high-frequency cues via Gaussian-blur residuals. We fuse them into an attention map and normalize it to $[0, 1]$:

$$\boldsymbol{A} = \mathrm{Normalize}\big(E + \lambda_{\mathrm{hf}} H\big), \tag{6}$$

where E denotes gradient magnitude, H the high-frequency residual, and λ_{hf} balances their contributions. This Structure-wise Attention Map highlights vertebral boundaries and texture-rich regions.

Pixel-wise Weight Map Based on the Structure-wise Attention Map, we compute a Pixel-wise Weight Map $\boldsymbol{w} \in \mathbb{R}^{H\times W}$ using a warm-up scheduler and an adaptive quantile threshold with a soft Gate Mask. A linear warm-up factor $r(t)$ progressively increases the weighting strength after iteration t_s. An adaptive threshold τ is computed as the q-quantile of $\boldsymbol{A}$ to suppress noisy responses. The final weights are defined as:

$$\boldsymbol{w} = 1 + \alpha_{\text{eff}}\, \boldsymbol{G} \odot \boldsymbol{A}, \tag{7}$$

where $\alpha_{\text{eff}} = \alpha r(t)$ and $\boldsymbol{G}$ is the soft gate mask derived from τ. This design emphasizes informative regions while maintaining stable optimization dynamics during early-stage training.

2.4 Loss Function

Given a sampled viewpoint, the unified renderer produces a Rendered Corrected View $\hat{\boldsymbol{I}}$, which is supervised by the ground-truth image $\boldsymbol{I}$. We optimize the Gaussian parameters using a hybrid loss that combines an SSIM term with a Weighted Charbonnier reconstruction term reweighted by the SPWM:

$$\mathcal{L} = (1 - \lambda_{\text{dssim}})\, \mathcal{L}_{w\text{-charb}} + \lambda_{\text{dssim}}\, (1 - \text{SSIM})\,, \tag{8}$$

$$\mathcal{L}_{w\text{-charb}} = \frac{\sum_p w(p)\, \rho(e(p))}{\sum_p w(p) + \delta}. \tag{9}$$

Here, $\rho(\cdot)$ denotes the Charbonnier penalty, $e(p)$ is the per-pixel residual, and $w(p)$ is the SPWM weight at pixel p. This hybrid formulation improves boundary sharpness, suppresses low-contrast artifacts, and enhances cross-view structural consistency while maintaining stable optimization.

3 Experimental Results

3.1 Dataset and System Implementation

Dataset We employ two 3D spine datasets for cross-modality multi-view synthesis. CTSpine3D [8, 20] contains over 600 CT volumes, and FeSpine3D contains 100 3D ultrasound volumes. For each volume, multi-view projections are rendered under a known cone-beam geometry using a GPU-accelerated differentiable DRR operator implemented with TIGRE [3]. To ensure a strict separation between training and evaluation views, 50 training projections are uniformly sampled over 180° at 3.6° intervals, whereas 50 testing projections are generated from the corresponding held-out midpoint angles. Specifically, FeSpine3D is derived from clinically acquired fetal spine ultrasound data collected under routine clinical protocols, and all data were de-identified and anonymized prior to use. For more details and usage instructions, please refer to the GitHub repository.

Implementation Details Our model is implemented in PyTorch and follows the training and densification–pruning pipeline of X-Gaussian [4]. All experiments are conducted on a single NVIDIA RTX 3090 GPU. We optimize the model with Adam ($\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 10^{-15}$) for 20,000 iterations,

Table 1. Quantitative comparison of reconstruction quality and efficiency on CTSpine3D and FeSpine3D datasets. Best results in **bold**, while second best underlined.

Dataset	Method	Inference Speed (fps)↑	Train Time (min)↓	PSNR↑	SSIM↑
CTSpine3D	IntraTomo [23]	0.27	33	35.27	0.9641
	NeRF [16]	0.02	410	37.61	0.9802
	TensoRF [6]	0.06	185	35.60	0.9600
	NAF [25]	0.37	93	36.64	0.9741
	Ours	**113.14**	**15**	**46.54**	**0.9938**
FeSpine3D	IntraTomo	0.17	29	23.60	0.8197
	NeRF	0.02	466	23.86	0.8286
	TensoRF	0.04	152	20.79	0.6838
	NAF	0.23	89	23.78	0.8271
	Ours	**143.17**	**8**	**40.82**	**0.9846**

using learning rates of 0.002, 0.008, 0.005, and 0.001 for feature parameters, opacity, scaling, and rotation, respectively. Densification is performed every 200 iterations from iteration 500 to 8,000 with a gradient threshold of 1.5×10^{-4}. For the proposed SPWM, we set $\alpha = 0.8$ and use a warm-up schedule that starts at iteration 2,000 and linearly increases over 5,000 iterations. The structure-prior attention map is constructed with $\lambda_{\text{hf}} = 0.5$ and a 3×3 Gaussian blur kernel with $\sigma = 1.5$, while adaptive gating uses a quantile threshold of $q = 0.85$.

Evaluation Metrics We evaluate UniSpine-GS in terms of both reconstruction quality and computational efficiency. For image quality, we report Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) on held-out views, following common practice in novel-view synthesis and sparse-view reconstruction [4, 6, 16, 25]. For efficiency, we measure inference speed in frames per second (fps) during rendering and report training time in minutes for comparison experiments and in seconds for ablation studies, all under the same hardware settings. All results are averaged over test views and cases. Higher is better for PSNR/SSIM/fps, and lower is better for training time.

3.2 Comparison Experiments

As shown in Table 1, UniSpine-GS achieves the best reconstruction quality and substantially higher efficiency on both CTSpine3D (X-ray) and FeSpine3D (ultrasound). On CTSpine3D, our method attains 46.54 PSNR and 0.9938 SSIM, surpassing the best-performing baseline (NeRF) by 8.93 dB in PSNR and 0.0136 in SSIM. UniSpine-GS also renders at 113.14 fps, achieving over 300× speedup compared with NAF (0.37 fps), and converges in 15 min, which is 6.2× faster than NAF (93 min) under the same setting.

On FeSpine3D, UniSpine-GS attains 40.82 PSNR and 0.9846 SSIM, surpassing the strongest baseline (NeRF) by 16.96 dB in PSNR and 0.1560 in SSIM. It runs at 143.17 fps, achieving an approximately 622× speedup over the fastest

Table 2. Ablation analysis of our model's components on CTSpine3D and FeSpine3D datasets. "✓" and "×" denote whether a component is enabled. Best results in **bold**, while second best underlined. We omit the "SPWM-only" setting because SPWM is designed as a loss reweighting module on top of a projection-consistent forward model and stable geometry initialization. Without RIRF and ACUI, the baseline optimization is often unstable and the resulting comparison is less informative.

Dataset	Method	RIRF +ACUI	SPWM	PSNR↑	SSIM↑	Train Time (s)↓
CTSpine3D	3DGS	×	×	37.83	0.9856	1132
	X-Gaussian	✓	×	<u>46.04</u>	<u>0.9926</u>	**756**
	UniSpine-GS	✓	✓	**46.54**	**0.9938**	<u>923</u>
FeSpine3D	3DGS	×	×	34.94	0.9734	850
	X-Gaussian	✓	×	<u>39.91</u>	<u>0.9796</u>	<u>782</u>
	UniSpine-GS	✓	✓	**40.82**	**0.9846**	**504**

baseline (NAF, 0.23 fps), and converges within just 8 min, which is about 3.6× faster than IntraTomo (29 min) and 58.3× faster than NeRF (466 min). Overall, these results consistently demonstrate that UniSpine-GS offers a superior quality–efficiency trade-off for multi-view spine reconstruction across both imaging modalities, enabling high-quality synthesis with fast training and inference.

3.3 Ablation Study

As summarized in Table 2, we analyze the contribution of each component on both CTSpine3D (X-ray) and FeSpine3D (ultrasound). On CTSpine3D, introducing the radiative design with RIRF and ACUI markedly improves performance over 3DGS, raising PSNR from 37.83 to 46.04 and SSIM from 0.9856 to 0.9926, while reducing training time from 1132 s to 756 s. This confirms the benefit of a projection-oriented forward model and geometry-aware initialization. Adding the proposed SPWM further improves fidelity, reaching 46.54 PSNR and 0.9938 SSIM. The slightly longer training time of 923 s is mainly due to extra per-iteration computation for constructing the weight map and the finer refinement it encourages near structural boundaries.

Consistent with the above findings, UniSpine-GS also achieves strong gains on FeSpine3D. From the 3DGS baseline, adding RIRF and ACUI boosts PSNR from 34.94 to 39.91 and SSIM from 0.9734 to 0.9796. With the SPWM, PSNR further rises to 40.82 and SSIM to 0.9846, while training time drops from 782 s to 504 s. This suggests that, unlike CT where SPWM emphasizes boundary refinement with slight computational overhead, in FeSpine3D it suppresses speckle-like noise and focuses optimization on structural regions, leading to faster convergence.

3.4 Visualization Analysis

Qualitative results on both CTSpine3D (X-ray) and FeSpine3D (ultrasound) are shown in Fig. 2, demonstrating that UniSpine-GS produces reconstructions closer to GT with sharper edges, richer mid-/high-frequency textures, and better local detail preservation. In the zoomed regions, competing methods tend to over-smooth vertebral boundaries and wash out subtle intensity variations, while UniSpine-GS maintains clearer structural contours and more consistent textures. Notably, in the ultrasound results of the competing methods, we observe low-contrast whitish artifacts, for example, in the lower region of the NeRF visualization. These artifacts partly explain their lower scores, whereas UniSpine-GS produces cleaner reconstructions with fewer such artifacts.

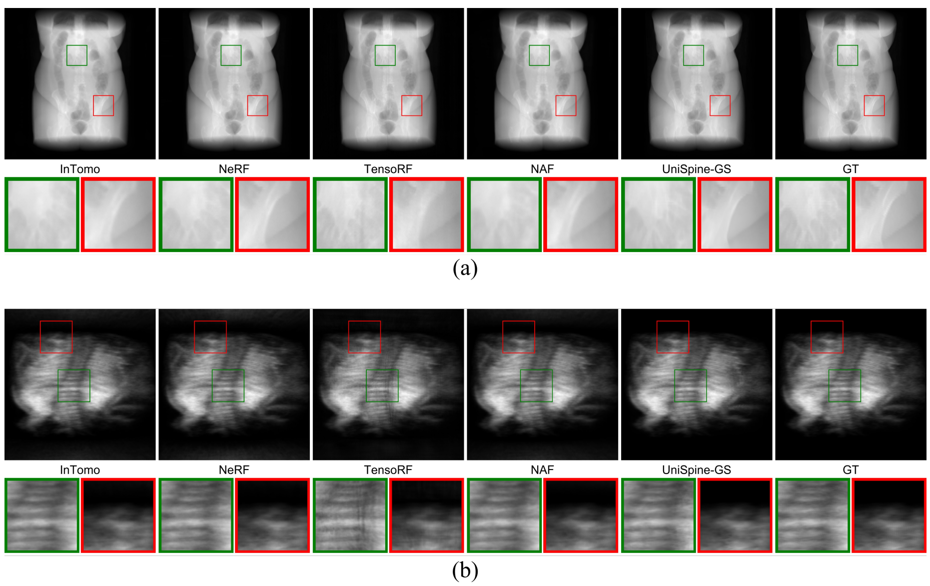

Fig. 2. Visualization results on (a) CTSpine3D and (b) FeSpine3D datasets.

4 Conclusion

We propose UniSpine-GS, an efficient physics-aware Gaussian framework for cross-modality multi-view spine image synthesis. With an explicit geometry-aware representation, a unified radiative forward operator, and a parameter-light SPWM, it enhances fine-grained fidelity and cross-view consistency while keeping training and inference fast. Experiments on CTSpine3D and the new FeSpine3D ultrasound dataset show consistent gains over neural field and radiative baselines, achieving a better quality–efficiency trade-off across modalities. Qualitative results demonstrate sharper boundaries and fewer low-contrast artifacts, especially in ultrasound. Future work will explore stronger anatomical priors and lightweight modality adapters, and will move toward more ultrasound-consistent differentiable forward models.

5 Acknowledgments

This work was supported in part by the National Key Research and Development Program of China under Grant 2023YFC2705702, in part by the National Natural Science Foundation of China under Grants 62225113 and U23B2048, in part by the Innovative Research Group Project of Hubei Province under Grants 2024AFA017, in part by the Science and Technology Major Project of Hubei Province under Grants 2024BAB046 and 2025BCB026, and in part by the New Cornerstone Science Foundation through the XPLORER PRIZE. This work was also supported by WHU-Kingsoft Joint Lab. The numerical calculations in this paper have been done on the supercomputing system in the Supercomputing Center of Wuhan University.

Disclosure of Interests

The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bao, Y., Ding, T., Huo, J., Liu, Y., Li, Y., Li, W., Gao, Y., Luo, J.: 3d gaussian splatting: Survey, technologies, challenges, and opportunities. IEEE Transactions on Circuits and Systems for Video Technology **35**(7), 6832–6852 (2025)
2. Barron, J.T., Mildenhall, B., Verbin, D., Srinivasan, P.P., Hedman, P.: Zip-nerf: Anti-aliased grid-based neural radiance fields. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 19697–19705 (2023)
3. Biguri, A., Dosanjh, M., Hancock, S., Soleimani, M.: Tigre: a matlab-gpu toolbox for cbct image reconstruction. Biomedical Physics & Engineering Express **2**(5), 055010 (2016)
4. Cai, Y., Liang, Y., Wang, J., Wang, A., Zhang, Y., Yang, X., Zhou, Z., Yuille, A.: Radiative gaussian splatting for efficient x-ray novel view synthesis. In: ECCV (2024)
5. Cai, Y., Wang, J., Yuille, A., Zhou, Z., Wang, A.: Structure-aware sparse-view x-ray 3d reconstruction. In: CVPR (2024)
6. Chen, A., Xu, Z., Geiger, A., Yu, J., Su, H.: Tensorf: Tensorial radiance fields. In: European conference on computer vision. pp. 333–350. Springer (2022)
7. Deng, T., Chen, Y., Zhang, L., Yang, J., Yuan, S., Liu, J., Wang, D., Wang, H., Chen, W.: Compact 3d gaussian splatting for dense visual slam. arXiv preprint arXiv:2403.11247 (2024)
8. Deng, Y., Wang, C., Hui, Y., Li, Q., Li, J., Luo, S., Sun, M., Quan, Q., Yang, S., Hao, Y., et al.: Ctspine1k: A large-scale dataset for spinal vertebrae segmentation in computed tomography. arXiv preprint arXiv:2105.14711 (2021)
9. Feng, C., Yu, W., Cheng, X., Tang, Z., Zhang, J., Yuan, L., Tian, Y.: Ae-nerf: Augmenting event-based neural radiance fields for non-ideal conditions and larger scenes. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 39, pp. 2924–2932 (2025)

10. Fu, Y., Liu, S., Kulkarni, A., Kautz, J., Efros, A.A., Wang, X.: Colmap-free 3d gaussian splatting. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 20796–20805 (2024)
11. Kazerouni, A., Aghdam, E.K., Heidari, M., Azad, R., Fayyaz, M., Hacihaliloglu, I., Merhof, D.: Diffusion models in medical imaging: A comprehensive survey. Medical image analysis **88**, 102846 (2023)
12. Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G., et al.: 3d gaussian splatting for real-time radiance field rendering. ACM Trans. Graph. **42**(4), 139–1 (2023)
13. Kumar, S., Awadhiya, B., Ratnakumar, R., Thalengala, A., Areeckal, A.S., Nanjappa, Y.: A review of 3d modalities used for the diagnosis of scoliosis. Tomography **10**(8), 1192–1204 (2024)
14. Liang, Z., Zhang, Q., Feng, Y., Shan, Y., Jia, K.: Gs-ir: 3d gaussian splatting for inverse rendering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 21644–21653 (June 2024)
15. Liu, D., Wang, Z., Chen, P.: Dsem-nerf: Multimodal feature fusion and global–local attention for enhanced 3d scene reconstruction. Information Fusion **115**, 102752 (2025)
16. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: Nerf: Representing scenes as neural radiance fields for view synthesis. Communications of the ACM **65**(1), 99–106 (2021)
17. Müller, T., Evans, A., Schied, C., Keller, A.: Instant neural graphics primitives with a multiresolution hash encoding. ACM transactions on graphics (TOG) **41**(4), 1–15 (2022)
18. Tang, J., Ren, J., Zhou, H., Liu, Z., Zeng, G.: Dreamgaussian: Generative gaussian splatting for efficient 3d content creation. arXiv preprint arXiv:2309.16653 (2023)
19. Wu, G., Yi, T., Fang, J., Xie, L., Zhang, X., Wei, W., Liu, W., Tian, Q., Wang, X.: 4d gaussian splatting for real-time dynamic scene rendering. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 20310–20320 (2024)
20. Yu, C., Huang, N., Su, H., Sun, C., Du, B.: Veganet: Conditional multi-view medical image correction via x-gaussian modeling for efficient perspective alignment. In: 2025 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 3177–3183. IEEE (2025)
21. Yu, W., Cai, Y., Zha, R., Fan, Z., Li, C., Yuan, Y.: X2-gaussian: 4d radiative gaussian splatting for continuous-time tomographic reconstruction. In: ICCV (2025)
22. Yugay, V., Li, Y., Gevers, T., Oswald, M.R.: Gaussian-slam: Photo-realistic dense slam with gaussian splatting. arXiv preprint arXiv:2312.10070 (2023)
23. Zang, G., Idoughi, R., Li, R., Wonka, P., Heidrich, W.: Intratomo: self-supervised learning-based tomography via sinogram synthesis and prediction. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1960–1970 (2021)
24. Zha, R., Lin, T.J., Cai, Y., Cao, J., Zhang, Y., Li, H.: R2-gaussian: Rectifying radiative gaussian splatting for tomographic reconstruction. In: NeurIPS (2024)
25. Zha, R., Zhang, Y., Li, H.: Naf: neural attenuation fields for sparse-view cbct reconstruction. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 442–452. Springer (2022)
26. Zhou, Y., Wu, C., Wang, X., Zheng, Y.: Uxdiff: Synthesis of x-ray image from ultrasound coronal image of spine with diffusion probabilistic network. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 3–12. Springer (2023)