
DDGS-CT: Direction-Disentangled Gaussian Splatting for Realistic Volume Rendering

Zhongpai Gao* Benjamin Planche* Meng Zheng
Xiao Chen Terrence Chen Ziyang Wu
United Imaging Intelligence, Boston, MA
{first.last}@uii-ai.com

Abstract

Digitally reconstructed radiographs (DRRs) are simulated 2D X-ray images generated from 3D CT volumes, widely used in preoperative settings but limited in intraoperative applications due to computational bottlenecks, especially for accurate but heavy physics-based Monte Carlo methods. While analytical DRR renderers offer greater efficiency, they overlook anisotropic X-ray image formation phenomena, such as Compton scattering. We present a novel approach that marries realistic physics-inspired X-ray simulation with efficient, differentiable DRR generation using 3D Gaussian splatting (3DGS). Our direction-disentangled 3DGS (DDGS) method separates the radiosity contribution into isotropic and direction-dependent components, approximating complex anisotropic interactions without intricate runtime simulations. Additionally, we adapt the 3DGS initialization to account for tomography data properties, enhancing accuracy and efficiency. Our method outperforms state-of-the-art techniques in image accuracy. Furthermore, our DDGS shows promise for intraoperative applications and inverse problems such as pose registration, delivering superior registration accuracy and runtime performance compared to analytical DRR methods.

1 Introduction

Motivation. Digitally reconstructed radiographs (DRRs) are simulated (*in silico*) 2D X-ray images rendered from 3D computational tomography (CT) volumes. While DRRs are widely utilized in preoperative settings, such as optimizing dose delivery and in radiation oncology [20, 41, 6], their potential intraoperative applications remain underexplored due to computational bottlenecks and naive modeling capability [1, 27]. For instance, real-time multimodal registration for image-guided procedures is currently impractical because of the time-consuming process of generating DRRs and integrating them into slice-to-volume registration [36].

The advent of GPU-accelerated computing has significantly improved the efficiency of radiography simulators [3, 39, 5], though they still fall short of meeting the stringent requirements of real-time applications. Recent developments in inverse graphics have led to the creation of differentiable DRR renderers (*e.g.*, DiffDRR [11], X-Gaussian [7], GaSpCT [25]), which show promise for inverse problems such as 2D/3D CT image registration [12]. However, these methods do not fully capture the complexities of X-ray image formation. They typically use computationally efficient ray-tracing techniques that model the attenuated photon fluence at each detector pixel by accumulating Hounsfield unit (HU) values along the 3D view-ray through the voxelized CT volume. This approach, while fast, cannot account for complex noise-inducing physical effects such as Compton scattering and beam hardening [14], crucial for accurate X-ray imaging.

*Equal contribution

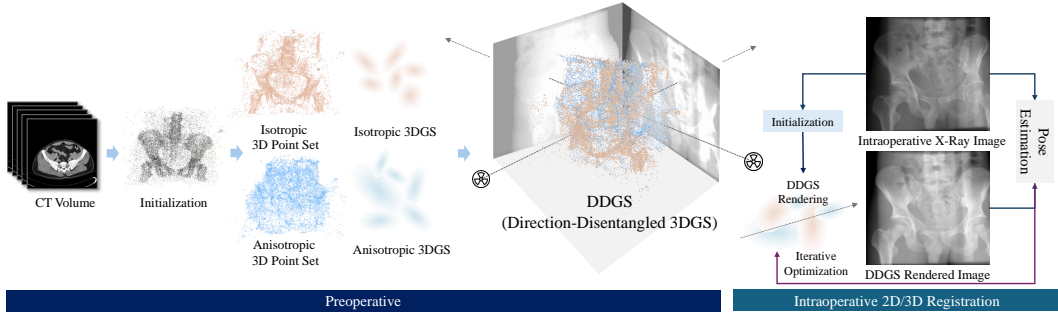


Figure 1: Proposed pipeline of Direction-Disentangled 3D Gaussian Splatting (DDGS).

Physics-based Monte-Carlo simulations [2, 15, 3, 4] do not suffer from these limitations. Relying on the decomposition of CT volumes into realistic materials via HU thresholding, they accurately simulate single-photon transport and probabilistically evaluate photon-matter interactions (photo-electric effect, Compton scattering, Rayleigh scattering, *etc.*) to determine the final attenuation and hit-region of each photon. This approach provides a more accurate representation of X-ray image formation; but the computational intensity of these methods—requiring $>10^8$ iterations to render an image—makes them impractical for real-time applications and inverse problems.

In this paper, we propose a novel, more elegant balance between (a) realistic physics-inspired X-ray simulation and (b) efficient, differentiable DRR generation, as illustrated in Figure 1.

Contributions. Our approach leverages 3D Gaussian splatting (3DGS) [17] for efficient DRR generation from CT volumes, similar to X-Gaussian [7] and GaSpCT [25]. However, we introduce a crucial adaptation: our direction-disentangled 3DGS (DDGS) method accounts for the dependencies of X-ray image formation on light directionality. Specifically, we decompose the radiosity contribution of 3D regions into two independent components: a non-directional component modeling the isotropic X-ray interactions with the corresponding material, and a direction-dependent component approximating higher-dimensional anisotropic interactions, such as scattering. Unlike the more generic view-dependent spherical-harmonics decomposition adopted by common 3DGS solutions [10], ours is tuned to the specific behaviors of high-energy photons. This formulation allows our method to implicitly learn and reproduce the effects of scattering on X-ray images without the need for complex photon-transport simulations at runtime.

Additionally, we adapt the initialization phase of 3DGS to account for the properties of tomography data. We adopt a geometry- and intensity-based dual strategy to sample 3D points that correspond to: (a) contact regions between different anatomical elements, where photon transport is likely to show discontinuities; and (b) homogeneous regions, where the primary attenuation contribution can be sparsely modeled but could benefit from view-dependent noise modeling.

We demonstrate that our solution achieves higher image accuracy compared to state-of-the-art methods across various medical benchmarks. Additionally, we showcase the efficient application of our method to inverse problems, such as pose registration, highlighting its potential for real-time intraoperative use cases. Our DDGS achieves better registration accuracy and runtime performance than analytical DRR methods.

2 Related Work

The generation of DRRs and the simulation of X-ray images have been subjects of extensive research due to their critical roles in medical imaging, particularly in radiotherapy planning and diagnostic imaging [20, 41, 6, 43]. In this section, we provide some insight on X-ray imaging, traditional simulation tools, and the recent integration of machine-learning techniques.

X-ray Image Simulation. X-ray imaging relies on the interactions of high-energy photons—emitted by the X-ray source of the scanner—with the matter (*e.g.*, body part) placed in between the X-ray source of the scanner and its radiation-sensitive detector cells (dexels). Imaging contrast arises from

differential absorption and scattering of X-ray photons by tissues, primarily through the photoelectric effect, Compton scattering, and, to a lesser extent, beam hardening for polychromatic scanners. The photoelectric effect involves the absorption of an X-ray photon by an atom and is more pronounced in higher atomic number materials like bone, causing them to appear whiter on X-ray images²; whereas Compton scattering occurs when an X-ray photon collides with an outer electron, resulting in the photon being deflected and losing energy, which contributes to image noise and reduced contrast [14, 29].

A variety of Monte Carlo (MC) simulation suites [2, 15, 3, 4] model these X-ray interactions and light transport. They trace the paths of a large number of individual photons as they undergo probabilistic interactions within the material. By simulating numerous photon trajectories, Monte Carlo methods account for the random nature of photoelectric absorption and Compton scattering, as well as secondary effects such as fluorescence and electron transport. These MC solutions produce highly realistic simulations of radiographic images, though at a high computational cost which makes them inadequate for most DRR use-cases. Furthermore, for these tools to render realistic DRRs from CT volumes, one must first pre-process the CT data to assign proper materials and their corresponding physical properties to each voxel, which is essential for realistic simulation. This step is often approximated by decomposing the volume into a small set of predefined materials according to user-provided HU ranges (e.g., “air” material assigned to voxels with HU value < -800 , “lung” assigned to voxels in range $] - 800, -200]$, “fat” for $] - 200, -100]$, etc.). This process can be time-consuming and prone to errors, affecting the fidelity of MC-based DRR generation [9].

Efficient and Differentiable DRR Generation. Analytical DRR rendering methods [39, 38, 5] have been developed in parallel to the aforementioned MC suites for real-time or near-real-time applications. These methods have a much lower computational footprint, as they do not simulate individual photon interactions but instead use mathematical models to directly calculate the primary paths and approximated attenuations of X-rays. The seminal work of Siddon [31] on computing line integrals through a discretized CT volume is still at the core of most recent DRR solutions, e.g., which proposed differentiable formulation of the rendering steps [11] or integrated neural networks to add realistic noise to the output [35]. However, the inherent approximations in these methods can affect their accuracy, particularly in modeling artifacts present in heterogeneous tissues or complex anatomical structures, e.g., caused by scattering (non-primary light contributions).

Recently, researchers [7, 25] have tried to apply 3D Gaussian splatting (3DGS) techniques [17] to DRR rendering, i.e., optimizing a cloud of 3D Gaussians to approximate voxel data, thereby enabling faster rendering with minimal accuracy loss. However, these early solutions overlooked the specific nuances of X-ray imaging, directly applying 3DGS methods designed for natural imaging [25] or oversimplifying the physical properties of X-ray attenuation [25], leading to suboptimal modeling.

In this paper, we propose a novel formulation of 3D Gaussian splatting, tuned to more efficiently approximate X-ray imaging (time- and quality-wise).

3 Methodology

3.1 Preliminaries

Gaussian Splatting. 3DGS [17] is a point-based rendering technique using 3D Gaussians to explicitly represent scenes in a more compact manner than volumetric representations. Each Gaussian $g_i = \{\mu_i, \Sigma_i, \alpha_i, f_i\}$ is defined by its mean position $\mu_i \in \mathbb{R}^3$, covariance matrix $\Sigma_i \in \mathbb{R}^{3 \times 3}$ (usually decomposed into separate scale and rotation parameters), opacity $\alpha_i \in \mathbb{R}$, and radiance properties $f_i \in \mathbb{R}^k$, where k depends on the light contribution model adopted. E.g., $k = (L + 1)^2 \times 3$ for anisotropic, view-dependent RGB radiance decomposed into spherical harmonics (SH) of degree L , a model commonly used for natural-imaging applications. Images are rendered by casting view-rays through each pixel p into the scene and alpha-blending the Gaussian contributions to the final ray color C , as:

$$C(p) = \sum_{j \in n} c_j \sigma_j \prod_{l=1}^{j-1} (1 - \sigma_l) \quad \text{with } \sigma_i = \alpha_i e^{-\frac{1}{2}(p - \hat{\mu}_i) \hat{\Sigma}_i^{-1} (p - \hat{\mu}_i)}, \quad (1)$$

²X-ray film is historically white, turning black when exposed to high-energy photon.

where n is the number of Gaussians considered, c_j is the radiance (computed from \mathbf{f}_j , *e.g.*, via SH) of the j th Gaussian on the ray, and $\hat{\boldsymbol{\mu}}_j$ and $\hat{\boldsymbol{\Sigma}}_j$ are the image-plane projections of $\boldsymbol{\mu}_j$ and $\boldsymbol{\Sigma}_j$.

Since the aforementioned rasterization process is differentiable w.r.t. to the Gaussian parameters, the 3DGS representation of a scene can be learned via gradient-descent, given a set of 2D scene observations \mathbf{I}_{GT} and their corresponding camera parameters. A combination of ℓ_1 and SSIM [40] loss functions are commonly-adopted as criterion for this iterative optimization process [17]:

$$\mathcal{L} = (1 - \lambda)\ell_1(\mathbf{I}_{\text{GT}}, \mathbf{I}) + \lambda \text{SSIM}(\mathbf{I}_{\text{GT}}, \mathbf{I}), \quad (2)$$

with \mathbf{I} the rendered images (based on provided camera parameters), and λ a loss-weighting hyper-parameter. The initial state is typically obtained by sampling relevant 3D points in the scene domain for the Gaussians, *e.g.*, from a structured-from-motion (SfM) point-cloud [34, 28].

Application to DRR. The volumetric rendering performed by 3DGS methods is conceptually similar to the one performed by analytical DRR methods [31, 11], *i.e.*, aggregating the attenuation values of light-rays cast through the voxelized CT volume. This similarity has recently motivated researchers [7, 25] to apply 3DGS models to DRR applications, driven by the compactness of 3DGS representations compared to voxel data (addressing the $\mathcal{O}(n^3)$ complexity inherent to voxel rendering). By optimizing a 3DGS model to approximate the visual properties of the CT volume, the resulting representation can render accurate DRRs much faster than traditional methods, making it more suitable for real-time *intra-operative* visualization. Consequently, the use of traditional, slower rendering solutions can be limited to generating DRR targets (\mathbf{I}_{GT}) for the offline (*i.e.*, *pre-operative*) training.

However, these early 3DGS-for-DRR solutions [7, 25] do not account for noise-inducing photon interactions (*e.g.*, scattering) when applying the analytical methods [30, 5] to render their training data. Since scatter-free DRRs are not affected by ray directions (*i.e.*, the attenuation at each point is naively considered isotropic), Cai et al. [7] simplified the radio-intensity function of each Gaussian to $c_i = \text{sigmoid}(\mathbf{b} \cdot \mathbf{f}_i)$, where $\mathbf{b} \in \mathbb{R}^k$ is a direction-independent optimizable basis vector shared by all splats. Furthermore, their Gaussian point-cloud is initialized via evenly-spaced Cai et al. [7] or random-uniform [25] sampling of the scanned 3D space, *i.e.*, ignoring relevant geometrical and material properties of the target CT data. Inadequate initialization strategies negatively impact the convergence of 3DGS models and the resulting image accuracy [16]. In this work, we propose to reformulate both the initialization strategy and the radiance function of the 3D Gaussians to address these limitations.

3.2 Disentanglement of Isotropic and Anisotropic 3D Gaussians

We argue that, while X-ray scattering models are incompatible with the 3DGS rendering formulation (secondary rays would have to be cast, increasing the computational footprint exponentially), their anisotropic impact on X-ray imaging can be approximated to some extent by direction-dependent radiance functions. However, modeling this high-dimensional residual contribution is only meaningful if it appears in the training DRRs, *i.e.*, if the target radiographs are rendered using physics-inspired simulation tools. This is not always possible, *e.g.*, if the mapping from HU values to materials for the target CT scans is not provided. Ideally, in such cases, when only scatter-free DRRs can be generated, the radiance function should gracefully degrade to a lighter anisotropic formulation (*e.g.*, similar to [7]), to avoid unnecessary computations that could impact the model optimization, as well as its runtime performance. Moreover, X-ray interactions with matter themselves involve both isotropic (*e.g.*, photoelectric absorption/fluorescence, Rayleigh scattering) and anisotropic (*e.g.*, Compton scattering) phenomena [14, 22]. Therefore, for the sake of both modularity and accuracy, we propose to decompose our DRR representation into isotropic Gaussians g_i^{iso} and anisotropic, direction-dependent ones g_j^{dir} . We formulate their respective radiosity functions as:

$$c_i^{\text{iso}} = \text{sigmoid}(\mathbf{b}^{\text{iso}} \cdot \mathbf{f}_i^{\text{iso}}) \quad ; \quad c_j^{\text{dir}} = \text{sigmoid}(Y_{1..L}(\theta, \phi) \cdot \mathbf{B}^{\text{dir}} \mathbf{f}_j^{\text{dir}}), \quad (3)$$

where $\mathbf{f}_i^{\text{iso}}, \mathbf{f}_j^{\text{dir}} \in \mathbb{R}^k$ are feature vectors respectively encoding the contribution of isotropic and anisotropic Gaussians, $\mathbf{b}^{\text{iso}} \in \mathbb{R}^k$ is a global isotropic basis vector (similar to [7]), and $\mathbf{B}^{\text{dir}} \in \mathbb{R}^{k_L \times k}$ is a second, anisotropic basis matrix that interacts with the L -degree spherical-decomposition $Y : \mathbb{R}^2 \mapsto \mathbb{R}^{k_L}$ applied to the ray angles θ, ϕ without constant term (degree = 0), resulting in $k_L = L(L + 2)$.

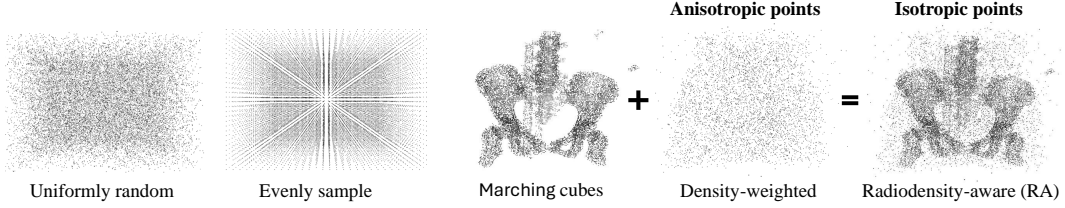


Figure 2: Illustration of the different sampling strategies for 3DGS initialization.

It is important to note that, while the usual spherical-harmonics representation for Gaussian splatting also contains isotropic (degree = 0) and anisotropic (degree > 0) components, the model assumes that both contributions are co-located (*i.e.*, they belong to the same 3D Gaussian g_i of position μ_i and covariance Σ_i). To account for the complexity of X-ray light transport, we relax this co-location constraint, *i.e.*, modeling isotropic and anisotropic contributions via distinct 3D Gaussians (g_i^{iso} and g_j^{dir}). We demonstrate in our evaluation that this disentanglement results in both higher image quality and lighter representation.

3.3 Initialization via Radiodensity-Aware Dual Sampling (RADS)

Our second main contribution targets the initialization of the point-cloud that supports the Gaussian-mixture modeling of the 3D data. Since 3DGS was defined for natural imaging (*e.g.*, ignoring sub-surface light transport), common initialization strategies limit their sampling to scene-surface points (SfM- [17] or depth-guided [24] subsampling). Such sampling of the 3D space is, however, inadequate for volumetric data, ignoring possibly salient regions. While uniform sampling of the CT space could be considered [7, 25], we argue that such a strategy is also suboptimal for anatomical data, discarding domain-relevant properties. We demonstrate empirically (*c.f.* Section 4) that uniform sampling results in slower convergence of the 3DGS representation and degraded runtime performance.

Therefore, we propose a novel twofold sampling strategy that accounts for the specific nature of CT data and for our dual isotropic/directional 3DGS model. Based on the radiodensity values contained in the target CT volumes, our solution samples two set of 3D points $\mathbf{P}^{\text{mc}} \in \mathbb{R}^{n_1 \times 3}$ and $\mathbf{P}^{\text{dw}} \in \mathbb{R}^{n_2 \times 3}$ with distinct, complementary distributions (n_1, n_2 scalar hyper-parameters).

Not unlike [24], the first set of points is sampled by applying the marching-cubes algorithm [21] to extract points at the interface between materials with different physical properties (*i.e.*, distinct HU responses in the CT volume). These interface regions typically need careful modeling when simulating X-ray imaging, as they result in discontinuous photon transport.

To complement this first set of points, the second set is semi-randomly sampled from the voxel centroids, according to a uniform distribution weighted by the voxels’ radiodensities. *I.e.*, voxels with higher radiodensity, contributing to the X-ray attenuation, are more likely to be picked to initialize the Gaussians. In other words, this radiodensity-weighted strategy identifies additional points within the homogeneous regions of the CT volume that also contributes to the imaging process.

Based on their distinct properties, we assign the entire \mathbf{P}^{mc} to initialize the isotropic Gaussians, whereas we equally split \mathbf{P}^{dw} to initialize both isotropic and direction-dependent Gaussians. All in all, our dual radiodensity-aware strategy improves the initialization of the 3DGS representation and is more explainable than prior work (see Figure 2).

4 Experiments

4.1 Experimental Protocol

Implementation. For evaluation, we set the decomposition degree to $L=1$, the feature dimension to $k=8$, and initial cloud sizes to $n_1=15,000$ and $n_2=10,000$. We apply a loss weight $\lambda = 0.2$. We choose the default threshold level (*i.e.*, the average of the minimum and maximum volume) for the marching-cubes algorithm. DDGS training is performed on a single NVIDIA A100 using an Adam

Table 1: Comparison of Gaussian splatting-based DRR rendering techniques, on 2 datasets.

		3DGS [17] ($L = 1$)			X-Gaussian [7] ($k = 32$)			DDGS (Ours , $L = 1, k = 8$)		
		# points↓	PSNR↑	SSIM↑	# points↓	PSNR↑	SSIM↑	# points↓	PSNR↑	SSIM↑
NAF-CT	abdomen	11,149	47.43	0.994	10,802	47.17	0.993	13,928	48.09	0.994
	chest	13,669	44.42	0.988	16,568	43.33	0.987	13,533	44.50	0.989
	foot	8,616	44.51	0.984	10,909	44.34	0.985	10,786	44.70	0.985
	jaw	17,902	40.47	0.973	22,318	40.02	0.972	19,665	40.57	0.974
	avg	12,834	44.21	0.985	15,149	43.72	0.984	14,478	44.47	0.986
CTPelvic1K-Dataset6	001	53,988	35.40	0.971	50,059	36.81	0.979	41,947	37.88	0.984
	002	49,099	37.03	0.982	48,113	37.79	0.986	43,933	38.43	0.988
	003	60,755	35.73	0.973	59,822	36.46	0.977	48,536	38.28	0.983
	004	42,349	38.87	0.982	37,243	39.84	0.985	39,176	40.35	0.986
	005	42,482	39.37	0.984	46,014	39.30	0.984	39,570	40.36	0.987
	006	53,832	37.14	0.983	57,197	37.42	0.983	47,398	38.26	0.986
	007	45,360	37.09	0.980	48,454	37.62	0.983	41,032	38.87	0.987
	008	51,211	38.03	0.980	45,750	38.70	0.982	43,577	39.81	0.986
	009	44,060	38.19	0.981	41,279	38.92	0.985	41,389	39.87	0.987
	010	38,691	37.80	0.984	40,030	37.55	0.985	39,691	38.73	0.987
	avg	48,183	37.47	0.980	47,396	38.04	0.983	42,625	39.08	0.986

optimizer [18] with a learning rate of 1.25×10^{-4} for \mathbf{b}^{iso} and \mathbf{B}^{dir} and 2.5×10^{-3} for \mathbf{f}^{iso} and \mathbf{f}^{dir} . Default 3DGS [17] learning rates are applied to the remaining parameters.

Datasets. We consider 4 datasets. (1) NAF-CT [42] includes four CT images of abdomen, chest, foot, and jaw. For each, we adopt TIGRE [5] to sample 50 evenly-distributed projections for training and 50 randomly-distributed ones for testing, in the range of $[-90^\circ, 90^\circ]$ (scatter-free DRRs). (2) CTPelvic1K [19] is a large dataset of pelvic CT images. We consider the first 10 scans of the sub-dataset6 in our experiments. We adopt DeepDRR [35] to generate 60 training (evenly-distributed) and 60 (randomly-distributed) testing DRRs (in the range of $[-60^\circ, 60^\circ]$), taking advantage of the scattering-modeling capability of DeepDRR to render realistic X-ray data. (3) Ljubljana [26] is a clinical dataset of 10 patients undergoing neurovascular surgery. Each patient underwent one CT and two X-ray angiography scans. The pose of each X-ray is also provided. Following DiffPose strategy [12], we randomly sample 900 projections for training and 100 for testing. (4) Finally, we provide some additional qualitative results on DeepFluoro [13], a collection of pelvic X-rays and CT images from six cadavers.

Metrics. We evaluate our method in two settings: novel-view synthesis and intraoperative 2D/3D image registration (*i.e.*, pose estimation). For novel-view synthesis, we use the standard PSNR and SSIM as metrics. For intraoperative 2D/3D image registration, we consider the error both in terms of rotation (angular distance) and translation (Euclidean distance). We further measure the clinically-relevant registration accuracy of key anatomical landmarks for Ljubljana scans [26] where they are provided. Following [12], we compute the target registration error (TRE) w.r.t. the positioning of these landmarks after registration.

4.2 Novel-View Synthesis

We first validate our contributions in terms of the quality of rendered DRRs and compactness of our dual Gaussian-mixture representation.

Comparative Evaluation. We compare to 3DGS [17] and X-Gaussian [7] on the NAF-CT [42] and CTPelvic1K [19] datasets for the novel-view synthesis. For each method, we adopt the hyper-parameters recommended by their respective authors (*e.g.*, $L = 1$ and 3 for normal 3DGS [17] and $k = 8$ and 32 for X-Gaussian [7]).

As shown in Table 1, for NAF-CT [42] (without scatter-related noise), our DDGS outperforms both 3DGS and X-Gaussian in terms of PSNR and SSIM. For CTPelvic1K [19] (with scatter), our DDGS significantly outperforms 3DGS and X-Gaussian in both PSNR and SSIM while using considerably

Table 2: Novel-view synthesis evaluation on CTPelvic1K data.

		Iteration=500		2000		7000		15000		30000	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DDGS (Ours) ($L = 1, k = 8$)	001	25.77	0.908	30.46	0.944	35.18	0.975	36.91	0.981	38.04	0.984
	002	27.64	0.921	31.60	0.954	35.33	0.977	36.90	0.984	38.30	0.987
DDGS (Ours) ($L = 3, k = 8$)	001	26.20	0.909	31.50	0.950	35.65	0.978	37.68	0.983	38.15	0.985
	002	27.87	0.923	32.23	0.956	36.02	0.979	37.87	0.986	38.97	0.989
3DGS [17] ($L = 1$)	001	23.56	0.873	27.52	0.914	31.66	0.951	33.64	0.963	35.36	0.970
	002	24.55	0.891	29.17	0.932	32.82	0.965	35.33	0.977	37.25	0.982
3DGS [17] ($L = 3$)	001	23.55	0.872	27.54	0.914	33.34	0.962	36.01	0.975	37.12	0.980
	002	24.55	0.891	28.96	0.931	34.50	0.973	37.72	0.985	38.75	0.988
X-Gaussian [7] ($k = 8$)	001	17.31	0.794	24.79	0.877	29.00	0.923	30.72	0.933	33.39	0.947
	002	13.45	0.742	27.19	0.904	32.72	0.962	34.52	0.973	36.04	0.978
X-Gaussian [7] ($k = 32$)	001	26.34	0.854	31.28	0.950	33.56	0.968	35.07	0.974	36.83	0.979
	002	28.32	0.868	33.44	0.966	35.85	0.980	36.55	0.984	37.89	0.986

Table 3: Ablation study w.r.t. the proposed disentangled isotropic/anisotropic representations.

		DDGS (Ours)		direct-entangled		direct-independent		direct-dependent		3DGS-disentangled	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
001		38.04	0.984	36.17	0.977	35.93	0.976	32.48	0.942	35.32	0.971
002		38.30	0.987	37.32	0.984	36.96	0.983	33.04	0.953	35.38	0.973

fewer points. Therefore, our DDGS method is more effective in simulating realistic X-ray images from CT scans.

Furthermore, we compare our DDGS (with degrees $L = 1$ and $L = 3$ and feature dimension $k = 8$) with traditional 3DGS (degrees $L = 1$ and $L = 3$) and X-Gaussian (feature dimensions $k = 8$ and $k = 32$) on the 001 and 002 data of CTPelvic1K [19] at iterations of 500, 2000, 7000, 15,000, and 30,000, as shown in Table 2. Our model with $L = 1$ and $k = 8$ performs comparably to 3DGS with $L = 3$ and outperforms X-Gaussian with $k = 32$ after 15,000 iterations. Experiments on more data can be found in Table S1. Our model with $L = 3$ achieves the best performance overall. Figure 3 presents a qualitative comparison of the testing set, illustrating both the synthetic views and the generated 3D Gaussian points.

It should also be highlighted that, even though both the original CT scan (as voxel grid) and our 3DGS-based solution are explicit representations, the latter is significantly more compact. *E.g.*, our model can represent a CTPelvic1K scan with only 42,625 Gaussians (*c.f.* Table 1), defined by 19 float values each (3D position/rotation + 3D covariance + opacity + feature vector) and a shared 32-dimensional basis vector \mathbf{b} , so 809,907 float values in total; whereas the original CT scans are each composed of $512 \times 512 \times 500 = 93,363,200$ values. Indeed, the voxel data may contain large homogeneous regions, which can be approximated with few Gaussians, hence a significant compression rate.

Finally, similar to [12], we qualitatively compare the results of recent DRR methods with actual X-ray images, for datasets providing both CT scans and real, posed projections (DeepFluoro [13], Ljubljana [26]). Results are shared in Figure 4, with our method achieving state-of-the-art PSNR. Further analyses are provided in Appendix B.

Ablation Study. Table 3 demonstrates the effectiveness of our direction-disentangled representation for 3D Gaussians. For several settings, we observe performance degradation when learning the direction-dependent and direction-independent components in the same 3D Gaussians (*direct-entangled*), learning the direction-independent 3D Gaussians alone (*direct-independent*), or learning the direction-dependent 3D Gaussians alone (*direct-dependent*). Additionally, we decouple the direction-dependent and direction-independent components in 3DGS [17] (*3DGS-disentangled*), which also showed lower performance compared to our learnable features. Figure 5 shows the rendered views from disentangled 3D Gaussians, highlighting that the direction-dependent 3D Gaussians

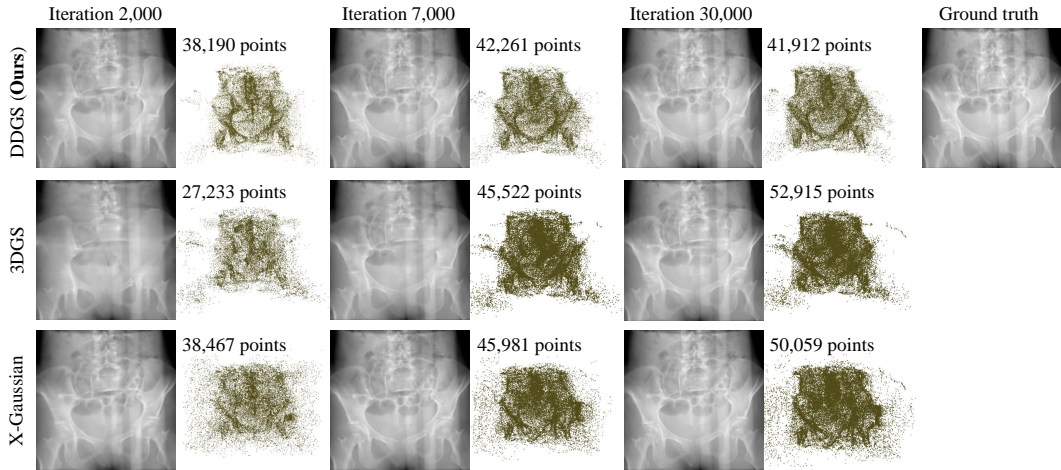


Figure 3: Visualization of the Gaussian cloud optimization for different methods.

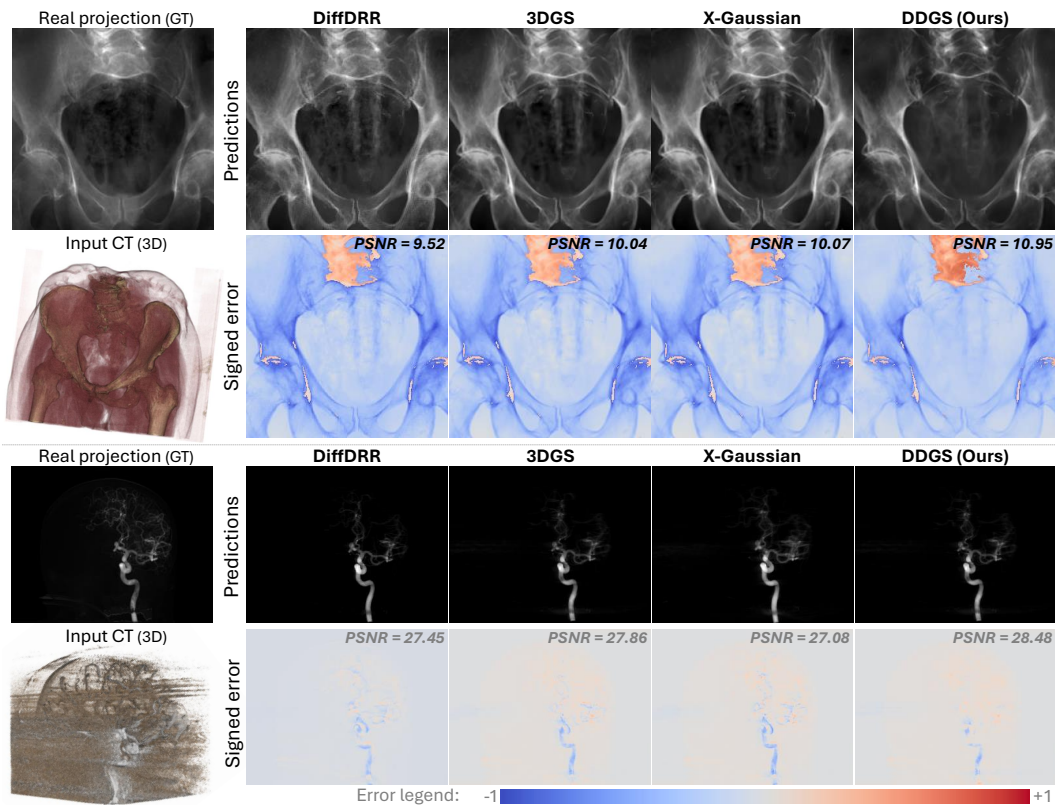


Figure 4: Qualitative comparison of DRRs and real scans from DeepFluoro and Ljubljana datasets.

focus more on bones or anatomical structures, while the direction-independent 3D Gaussians focus more on the background.

Table 4 shows the effectiveness of our proposed initialization scheme (radiodensity-aware) compared to other initialization methods. Our scheme outperforms both the uniformly random sampling used in 3DGS [17] and the even sampling method used in X-Gaussian [7].

At last, we conduct experiments on the impact of feature dimension k , as shown in Table 5. Our results indicate that increasing the feature dimension generally improves overall performance. However, the improvement becomes marginal once the feature dimension exceeds $k = 16$.

Table 4: Impact of initialization strategies on image quality during training, *i.e.*, measured on CTPelvic1K scans w.r.t. training iterations.

Iteration	Strategy	001	002	003	004	005	006	007	008	009	010	avg
500	RADS (ours)	25.86	27.72	25.24	28.52	30.52	28.31	29.23	27.92	28.95	29.20	28.15
	Random	25.29	26.73	24.66	27.40	28.46	26.35	27.23	26.43	27.34	28.53	26.84
	Even	24.96	26.64	24.40	26.60	27.54	26.63	25.06	26.01	26.15	26.92	26.09
7,000	RADS (ours)	35.02	37.06	36.47	36.49	36.79	34.52	35.56	36.03	36.44	35.55	35.99
	Random	34.39	35.78	34.18	35.83	36.37	34.36	35.20	35.74	36.30	35.26	35.34
	Even	34.22	36.07	33.65	36.31	36.64	34.13	34.97	35.85	35.61	35.33	35.28
30,000	RADS (ours)	37.88	38.43	38.28	40.35	40.36	38.26	38.87	39.81	39.87	38.73	39.08
	Random	37.17	38.36	38.15	40.09	39.97	38.36	38.69	39.37	39.85	38.84	38.89
	Even	37.33	38.51	37.83	39.58	39.12	38.27	38.87	39.88	39.46	38.95	38.78

Table 5: Impact of k dimensionality on image quality, measured on 2 CTPelvic1K scans (001, 002).

	$k = 1$		$k = 4$		$k = 8$		$k = 16$		$k = 32$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
001	34.60	0.965	36.81	0.979	38.04	0.984	38.41	0.986	38.27	0.987
002	33.66	0.964	37.08	0.982	38.30	0.987	38.76	0.989	38.89	0.989

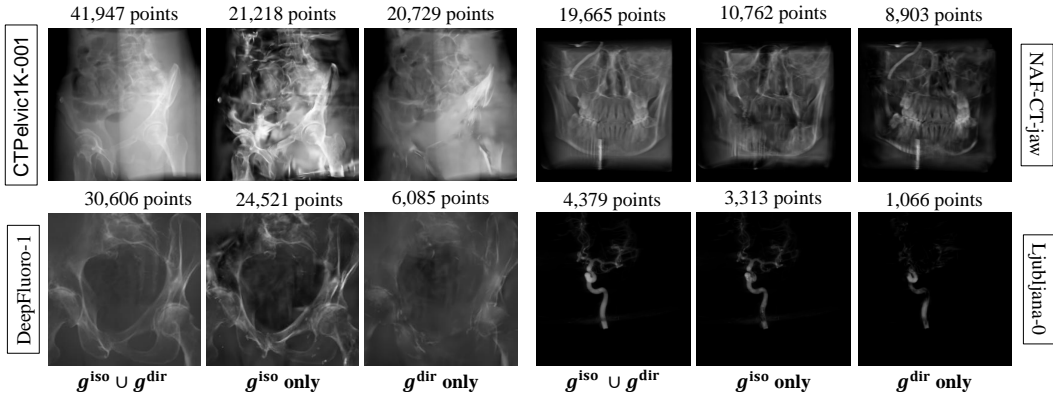


Figure 5: Visualization of the isotropic and anisotropic X-ray contributions to the rendered DRRs.

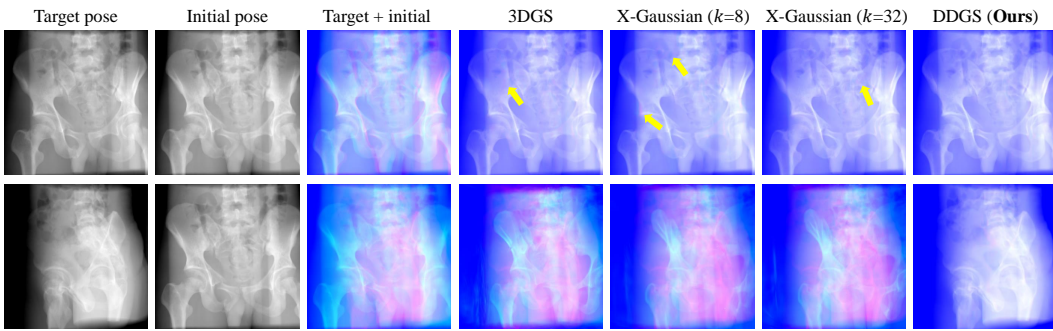


Figure 6: Illustration of the pose-registration convergence using various differentiable DRR renderers.

Table 6: Accuracy of image registration on the CTPelvic1K [19] dataset.

	3DGS [17]			X-Gaussian [7] ($k=8$)			X-Gaussian [7] ($k=32$)			DDGS (Ours)		
	001	002	003	001	002	003	001	002	003	001	002	003
Rot. error (deg)	0.134	0.062	0.229	0.326	0.163	0.312	0.097	0.070	0.266	0.067	0.067	0.152
Trans. error (mm)	1.788	1.163	3.108	4.037	2.144	3.781	1.378	1.288	3.483	1.285	1.092	2.180

Table 7: Registration evaluation w.r.t. anatomical landmark accuracy and time efficiency, on Ljubljana.

Method	DRR render time (ms) ↓			total optimization time (s) ↓			TRE (mm) ↓		
	mean	med	std	mean	med	std	mean	med	std
DiffDRR [11]	30.88	31.51	3.25	431.98	316.54	261.70	2.14	0.72	5.40
DDGS	7.46	6.94	2.67	121.45	39.93	117.67	0.80	0.65	0.51

4.3 Application to Downstream Task (2D/3D CT Image Registration)

We validate our solution in terms of applicability to downstream tasks, considering here the registration of intraoperative 2D X-ray images to preoperative 3D CT volumes; a task crucial to a variety of clinical applications.

We choose the first five testing images on the 001, 002, and 003 CT data of CTPelvic1K [19] for image registration. Considering this task as an optimization-based problem that can be solved via inverse graphics (*i.e.*, rendering synthetic images based on predicted poses and comparing to the target image, then backpropagating the difference to improve the pose prediction), we adopt the framework from iComMA [32], replacing their image rendered by DDGS. We use Adam optimizer with a learning rate of 0.05. We use the isocenter pose as the initialization pose. As shown in Table 6 and Figure 6, our DDGS achieves the lowest rotation and translation errors.

Finally, we also consider the experimental protocol proposed in DiffPose [12], where the authors rely on a pretrained pose-regression CNN to get a first rough estimation of the scanner pose, before refining said pose via a gradient-descent based optimization, leveraging their differentiable DRR renderer [11]. We adopt their pose-initialization network and testbed for the Ljubljana dataset, replacing their DiffDRR analytical renderer [11] by our DDGS, and compare to them in terms of runtime performance and landmark-registration accuracy. The results, reported in Table 7, highlight the convergence boost brought by our lighter solution.

5 Conclusion

We proposed DDGS, a novel 3DGS-based representation for efficient and realistic DRR generation from CT volumes. We extended traditional 3DGS by considering light directionality, crucial for accurate X-ray image modeling. By decomposing radiosity into non-directional and direction-dependent components, DDGS captures isotropic interactions and anisotropic effects. Unlike generic view-dependent spherical-harmonics decomposition, DDGS implicitly learns scattering effects, substantially improving the speed and accuracy of downstream clinical applications such as intraoperative 2D/3D registration.

Limitations. It should be noted that, similar to previous methods, our model expects an offline DRR renderer to produce realistic ground truth to train DDGS, which can be a limitation for some scenarios (*e.g.*, when such renderer or certain parameters required for realistic rendering are not available). Our direction-dependent function relating to anisotropic X-ray effects may also fail to approximate complex multi-bounce light scenarios. Recent developments in N -dimensional Gaussian splatting [8] show promising potential for modeling such cases. We should also note that the quality of DRRs is bound to the quality of the input CT scans. This issue affects both DRR methods and Monte-Carlo simulations (as the latter relies on the segmentation of CT scans into materials-specific regions – if a 3D scan is noisy, so will the material segmentation and simulation results). When our method uses simulators to get target images, it will face similar noise issues, though, our experiments show that DDGS barely introduces additional noise. Compensating for CT noise is an interesting, under-explored research direction that could benefit all DRR methods; but, we believe it is beyond the scope of our study.

Societal Impact. The above limitations may restrict the adoption of DDGS to specific use cases, and our method should undergo a clinical evaluation of the generated images in terms of anatomical accuracy. Nonetheless, we believe that our proposed method can positively impact both the computer vision and medical communities by providing a more efficient and versatile DRR tool.

References

- [1] John R Adler Jr, Martin J Murphy, Steven D Chang, and Steven L Hancock. Image-guided robotic radiosurgery. *Neurosurgery*, 44(6):1299–1306, 1999.
- [2] Sea Agostinelli, John Allison, K al Amako, John Apostolakis, H Araujo, Pedro Arce, Makoto Asai, D Axen, Swagato Banerjee, GJNI Barrand, et al. Geant4—a simulation toolkit. *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.
- [3] Andreu Badal and Aldo Badano. Accelerating monte carlo simulations of photon transport in a voxelized geometry using a massively parallel graphics processing unit. *Medical physics*, 36(11):4878–4880, 2009.
- [4] Julien Bert, Hector Perez-Ponce, Ziad El Bitar, Sébastien Jan, Yannick Boursier, Damien Vintache, Alain Bonissent, Christian Morel, David Brasse, and Dimitris Visvikis. Geant4-based monte carlo simulations on gpu for medical applications. *Physics in Medicine & Biology*, 58(16):5593, 2013.
- [5] Ander Biguri, Manjit Dosanjh, Steven Hancock, and Manuchehr Soleimani. Tigre: a matlab-gpu toolbox for cbct image reconstruction. *Biomedical Physics & Engineering Express*, 2(5):055010, 2016.
- [6] Marc A Bollet, Helen A McNair, Vibeke N Hansen, Andrew Norman, Una O’Doherty, Helen Taylor, Mark Rose, Rahul Mukherjee, and Robert Huddart. Can digitally reconstructed radiographs (drrs) replace simulation films in prostate cancer conformal radiotherapy? *International Journal of Radiation Oncology* Biology* Physics*, 57(4):1122–1130, 2003.
- [7] Yuanhao Cai, Yixun Liang, Jiahao Wang, Angtian Wang, Yulun Zhang, Xiaokang Yang, Zongwei Zhou, and Alan Yuille. Radiative gaussian splatting for efficient x-ray novel view synthesis. In *European Conference on Computer Vision*, pages 283–299. Springer, 2025.
- [8] Stavros Diolatzis, Tobias Zirr, Alexandr Kuznetsov, Georgios Kopanas, and Anton Kaplanyan. N-dimensional gaussians for fitting of high dimensional functions. *SIGGRAPH*, 2024.
- [9] O Dorgham, MH Ryalat, and M Abu Naser. Automatic body segmentation for accelerated rendering of digitally reconstructed radiograph images. *Informatics in Medicine Unlocked*, 20:100375, 2020.
- [10] Sara Fridovich-Keil, Alex Yu, Matthew Tancik, Qinhong Chen, Benjamin Recht, and Angjoo Kanazawa. Plenoxels: Radiance fields without neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5501–5510, 2022.
- [11] Vivek Gopalakrishnan and Polina Golland. Fast auto-differentiable digitally reconstructed radiographs for solving inverse problems in intraoperative imaging. In *Workshop on Clinical Image-Based Procedures*, pages 1–11. Springer, 2022.
- [12] Vivek Gopalakrishnan, Neel Dey, and Polina Golland. Intraoperative 2d/3d image registration via differentiable x-ray rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11662–11672, 2024.
- [13] Robert B Grupp, Mathias Unberath, Cong Gao, Rachel A Hegeman, Ryan J Murphy, Clayton P Alexander, Yoshito Otake, Benjamin A McArthur, Mehran Armand, and Russell H Taylor. Automatic annotation of hip anatomy in fluoroscopy for robust and efficient 2d/3d registration. *International journal of computer assisted radiology and surgery*, 15:759–769, 2020.
- [14] Bruce H Hasegawa. The physics of medical x-ray imaging. 1990.
- [15] Sébastien Jan, G Santin, D Strul, Steven Staelens, K Assié, D Autret, S Avner, R Barbier, M Bardies, PM Bloomfield, et al. Gate: a simulation toolkit for pet and spect. *Physics in Medicine & Biology*, 49(19):4543, 2004.
- [16] Jaewoo Jung, Jisang Han, Honggyu An, Jiwon Kang, Seonghoon Park, and Seungryong Kim. Relaxing accurate initialization constraint for 3d gaussian splatting. *arXiv preprint arXiv:2403.09413*, 2024.

- [17] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4):1–14, 2023.
- [18] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [19] Pengbo Liu, Hu Han, Yuanqi Du, Heqin Zhu, Yinhao Li, Feng Gu, Honghu Xiao, Jun Li, Chunpeng Zhao, Li Xiao, et al. Deep learning to segment pelvic bones: large-scale ct datasets and baseline models. *International Journal of Computer Assisted Radiology and Surgery*, 16: 749–756, 2021.
- [20] Frank Lohr, Oliver Schramm, Peter Schraube, Gabriele Sroka-Perez, Steffen Seeber, Gerd Schleppe, Wolfgang Schlegel, and Michael Wannemacher. Simulation of 3d-treatment plans in head and neck tumors aided by matching of digitally reconstructed radiographs (drr) and on-line distortion corrected simulator images. *Radiotherapy and oncology*, 45(2):199–207, 1997.
- [21] William E Lorensen and Harvey E Cline. Marching cubes: A high resolution 3d surface construction algorithm. In *Seminal graphics: pioneering efforts that shaped the field*, pages 347–353. 1998.
- [22] Alessandro Migliori. X-ray fluorescence analysis. <http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm>. Accessed: 2024-05-12.
- [23] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106, 2021.
- [24] Michael Niemeyer, Fabian Manhardt, Marie-Julie Rakotosaona, Michael Oechsle, Daniel Duckworth, Rama Gosula, Keisuke Tateno, John Bates, Dominik Kaeser, and Federico Tombari. Radsplat: Radiance field-informed gaussian splatting for robust real-time rendering with 900+ fps. *arXiv preprint arXiv:2403.13806*, 2024.
- [25] Emmanouil Nikolakakis, Utkarsh Gupta, Jonathan Vengosh, Justin Bui, and Razvan Marinescu. Gaspct: Gaussian splatting for novel ct projection view synthesis. *arXiv preprint arXiv:2404.03126*, 2024.
- [26] Franjo Pernus et al. 3d-2d registration of cerebral angiograms: A method and evaluation on clinical images. *IEEE transactions on medical imaging*, 32(8):1550–1563, 2013.
- [27] Terry M Peters. Image-guided surgery: from x-rays to virtual reality. *Computer methods in biomechanics and biomedical engineering*, 4(1):27–57, 2001.
- [28] Johannes L Schonberger and Jan-Michael Frahm. Structure-from-motion revisited. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4104–4113, 2016.
- [29] J Anthony Seibert and John M Boone. X-ray imaging physics for nuclear medicine technologists. part 2: X-ray interactions and image formation. *Journal of nuclear medicine technology*, 33(1): 3–18, 2005.
- [30] Gregory C Sharp, Rui Li, John Wolfgang, G Chen, Marta Peroni, Maria Francesca Spadea, Shinichiro Mori, Junan Zhang, James Shackelford, and Nagarajan Kandasamy. Plastimatch: an open source software suite for radiotherapy image processing. In *Proceedings of the XVIth International Conference on the use of Computers in Radiotherapy (ICCR), Amsterdam, Netherlands*, volume 3, 2010.
- [31] Robert L Siddon. Fast calculation of the exact radiological path for a three-dimensional ct array. *Medical physics*, 12(2):252–255, 1985.

- [32] Yuan Sun, Xuan Wang, Yunfan Zhang, Jie Zhang, Caigui Jiang, Yu Guo, and Fei Wang. icomma: Inverting 3d gaussians splatting for camera pose estimation via comparing and matching. *arXiv preprint arXiv:2312.09031*, 2023.
- [33] Andrea Tagliasacchi and Ben Mildenhall. Volume rendering digest (for nerf). *arXiv preprint arXiv:2209.02417*, 2022.
- [34] Shimon Ullman. The interpretation of structure from motion. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 203(1153):405–426, 1979.
- [35] Mathias Unberath, Jan-Nico Zaeck, Sing Chun Lee, Bastian Bier, Javad Fotouhi, Mehran Armand, and Nassir Navab. Deepdr—a catalyst for machine learning in fluoroscopy-guided procedures. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part IV 11*, pages 98–106. Springer, 2018.
- [36] IMJ Van Der Bom, Stefan Klein, Marius Staring, R Homan, Lambertus W Bartels, and Josien PW Pluim. Evaluation of optimization methods for intensity-based 2d-3d registration in x-ray guided interventions. In *Medical Imaging 2011: Image Processing*, volume 7962, pages 657–671. SPIE, 2011.
- [37] Delio Vicini, Wenzel Jakob, and Anton Kaplanyan. A non-exponential transmittance model for volumetric scene representations. *ACM Transactions on Graphics (TOG)*, 40(4):1–16, 2021.
- [38] Franck P Vidal and Pierre-Frédéric Villard. Development and validation of real-time simulation of x-ray imaging with respiratory motion. *Computerized Medical Imaging and Graphics*, 49: 1–15, 2016.
- [39] Franck Patrick Vidal, Manuel Garnier, Nicolas Freud, Jean-Michel Létang, and Nigel W John. Simulation of x-ray attenuation on the gpu. In *TPCG*, pages 25–32, 2009.
- [40] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4): 600–612, 2004.
- [41] Charles Yang, Michael Guiney, Peter Hughes, Samuel Leung, Kuen Hoe Liew, James Matar, and George Quong. Use of digitally reconstructed radiographs in radiotherapy treatment planning and verification. *Australasian Radiology*, 44(4):439–443, 2000.
- [42] Ruyi Zha, Yanhao Zhang, and Hongdong Li. Naf: Neural attenuation fields for sparse-view cbct reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 442–452. Springer, 2022.
- [43] Xinyang Zhang, Pengbo He, Yazhou Li, Xinguo Liu, Yuanyuan Ma, Guosheng Shen, Zhongying Dai, Hui Zhang, Weiqiang Chen, and Qiang Li. Dr-only carbon-ion radiotherapy treatment planning via deep learning. *Physica Medica*, 100:120–128, 2022.

Supplementary Material

In this supplementary material, we provide further methodological context and showcase additional quantitative and qualitative results to highlight further the contributions claimed in the paper.

A Methodological Insight

A.1 Transmittance in Natural vs. X-Ray Imaging

The exponential transmittance model $T(t)$ used in NeRF and Gaussian-splatting (GS) solutions for natural scenes—to describe the light attenuation as it travels through a medium from point $r(t_0)$ to $r(t)$ —is based on Beer-Lambert law [16, 23, 37, 33]:

$$T(t) = \exp\left(-\int_{t_0}^t a(r(s))\delta s\right), \quad (4)$$

where $r(t) = o+td$ is a 3D point at step t on a ray of origin o and direction d , and $a(p)$ is the medium’s absorption/extinction (linked to its density) at point p . Assuming a piece-wise homogeneous medium (with a_i the absorption of homogeneous segment l of length δ_l), this transmittance can be rewritten:

$$T(t) = \exp\left(-\sum_{l=1}^{l_t} a_l\delta_l\right) = \prod_{l=1}^{l_t} (1 - \sigma_l), \quad (5)$$

as in our Equation 3.1; with $\sigma_l = 1 - \exp(-a_l\delta_n)$ and l_t the end segment containing t . See [33] for proof.

X-ray attenuation from photoelectric effect (PE) follows the same model: at each step through the medium (*e.g.*, voxel), the probability of photon extinction (*i.e.*, averaged attenuation) is proportional to the atomic density. *E.g.*, if a ray travels through only 2 voxels, and if it is attenuated by $a_1 = 50\%$ at step 1 then $a_2 = 50\%$ again at step 2, then its total attenuation is 75% (*c.f.* $(100 - (100 - 50)^2)\%$), in accordance to Beer-Lambert.

Hence, Equation 3.1 (3DGS rendering) could be directly applied to DDR synthesis (assuming standard *neglog* scaling of absorption values into pixel ones [12, 35]), if we were to adopt a simplified model of X-ray imaging which only considers isotropic absorption (*i.e.*, following the Beer-Lambert law) and optionally omit the term $c_j \in \mathbb{R}^K$ (view-dependent radiance). This is the simplified model proposed in prior work [7].

In our paper, we propose to go beyond the isotropic simplification in X-Gaussian [7] and further account for anisotropic scattering of X-rays. This is why we preserve c_j and decompose it into two terms: an isotropic term c^{iso} and direction-dependent non-linear term c^{dir} (Equation 3.2). Note that we also fix the dimension K of these variables (*i.e.*, the number of output channels) to 1 (monochromatic DRR) instead of three in natural 3DGS (RGB).

A.2 Joint Gaussian Rasterization

As explained in Section 3, we define two distinct functions for the absorption contribution of the two Gaussian sets: one function is isotropic to approximate average radio-absorption (c_i^{iso}), and one function is direction-dependent to approximate the contribution of anisotropic Compton scattering to the image (c_j^{dir}), *c.f.* Equation 3.2.

These two distinct Gaussian sets could be rasterized separately, as done in Figure 5 as qualitative ablation. However, the two sets *need to be rendered together to ensure correct simulation*. Otherwise, Gaussians from one set occluded by others belonging to the other set could incorrectly contribute to the ray absorption. We refer the readers to Figure S1 for an explanatory diagram. There, point g_2^{dir} should not contribute to the final pixel value, as it is occluded by g_2^{iso} (*i.e.*, it absorbs all the remaining energy before the ray can reach g_2^{dir}). This could not be properly simulated if each Gaussian set is rasterized separately.

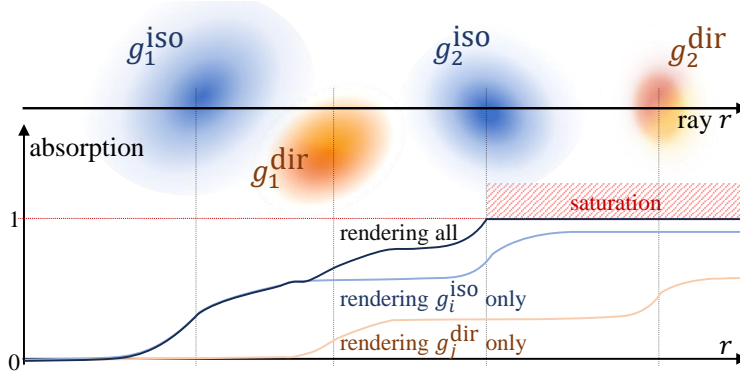


Figure S1: Importance of joint rasterization of Gaussian sets (simplified diagram).

By having 2 sets of Gaussians with distinct yet complementary contribution functions (isotropic versus direction-dependent functions), our solution can better model DRR imaging. During optimization, each Gaussian set will approximate its respective imaging effect, conditioned by its respective function.

We randomly split \mathbf{P}^{mc} (3D points sampled based on the CT scan distribution) into 2 sets, each subset contributing to the initialization positions of g_i^{iso} and g_j^{dir} (half the 3D points in \mathbf{P}^{mc} will serve as initialization for g_i^{iso} , along with the entire \mathbf{P}^{mc} set; and the other half of \mathbf{P}^{mc} will serve as initialization for g_j^{dir}). After initialization, during 3DGS-based model optimization, the points in each set of Gaussians can move/split/drop independently (along with the optimization of their respective covariance/opacity/feature values). Since the optimization of each set (isotropic and anisotropic) is conditioned on their respective differentiable absorption function, they will acquire distinct properties.

B Additional Experiments

B.1 Additional Comparison and Ablation Results

To further support the results in Table 2 and Table 3, we evaluate our method on all the 10 selected data from the CTPelvic1K dataset, as shown in Table S1.

Table S1: Novel-view synthesis evaluation on CTPelvic1K data.

	DDGS (ours) ($L = 1, k = 8$)		3DGS ($L = 1$)		X-Gaussian ($k = 8$)		DDGS (ours) ($L = 3, k = 8$)		3DGS ($L = 3$)		X-Gaussian ($k = 32$)		direct-entangled <i>c.f.</i> Ablation Study	
	PNSR	SSIM	PNSR	SSIM	PNSR	SSIM	PNSR	SSIM	PNSR	SSIM	PNSR	SSIM	PNSR	SSIM
001	38.04	.984	35.36	.970	33.39	.947	38.15	.985	37.12	.980	36.83	.979	36.17	.977
002	38.30	.987	37.25	.982	36.04	.978	38.97	.989	38.75	.988	37.89	.986	37.32	.984
003	38.21	.983	35.81	.974	33.99	.965	39.42	.986	39.09	.984	35.84	.976	35.32	.975
004	40.54	.987	38.43	.980	37.18	.974	40.74	.987	40.69	.986	39.06	.984	39.70	.985
005	39.88	.987	39.44	.985	37.50	.978	40.02	.987	40.68	.989	39.24	.985	39.36	.986
006	38.12	.986	36.69	.982	35.03	.977	38.79	.986	38.65	.987	37.14	.983	36.89	.983
007	38.91	.987	36.88	.978	35.51	.970	39.37	.988	38.58	.984	38.10	.987	37.72	.985
008	39.83	.982	37.85	.979	36.70	.973	40.19	.986	39.70	.986	38.57	.982	38.53	.982
009	39.67	.987	37.94	.980	37.60	.979	40.35	.988	39.16	.983	38.99	.987	38.74	.985
010	38.60	.987	37.50	.982	36.30	.978	39.18	.988	39.13	.986	37.89	.986	37.20	.985
avg	39.01	.986	37.32	.979	35.92	.972	39.52	.987	39.16	.985	37.96	.984	37.70	.983

B.2 Comparison to Real X-ray Projections

The main purpose of DRRs is to provide quick visualization to clinicians (with the key metrics being speed and visibility), as well as to integrate larger imaging applications (with priority given again to speed and feature-level similarity w.r.t. real data). Therefore, prior DRR papers (traditional [11, 30, 31] or GS-based [7, 25]) mostly evaluate on downstream tasks (*e.g.*, pose registration). A few evaluate the image quality compared to other DRR tools (DiffDRR [11] compares to Siddon's

Table S2: Quantitative comparison of analytical DRRs to real images from Ljubljana dataset.

Methods	SSIM			PSNR		
	mean [↑]	med [↑]	std [↓]	mean [↑]	med [↑]	std [↓]
DiffDRR [11]	.404	.440	.179	24.16	24.35	2.82
3DGS [17]	.409	.518	.199	22.57	23.44	2.78
X-Gaussian [7]	.409	.521	.198	22.58	23.42	2.78
DDGS (ours)	.459	.558	.204	25.61	27.35	3.48

and Plastimatch [30]; [7] to other NVS baselines). We found that only DiffPose (based on DiffDRR) [12] provides a qualitative comparison to some real scans.

Quantitative comparison to real data is challenging. Not all imaging parameters are usually provided or accurate enough to create matching DRRs (w.r.t. intensity, CT pose, *etc.*). It is especially hard to render DRRs aligned with real scans with existing Monte-Carlo (MC) simulation tools [2–4, 15]; which is likely why *no DRR paper has performed such evaluation*. Their interfaces and custom coordinate conventions are not compatible with pose annotations in public CT/X-ray datasets, and their documentation/support is lacking. Bridging the convention gap between MC and analytical communities would greatly benefit our domain and could be the focus of our next effort.

Nevertheless, we do provide an indirect comparison to real projections. *E.g.*, the registration experiments (Table 6) implicitly provide insight into the similarity between real and synthetic data, as pose optimization is done by comparing DRRs to real target scans.

In this supplementary, we share a more direct comparison. We use our DDGS-based registration method to refine GT poses in Ljubljana data (*c.f.* aforementioned GT inaccuracies), then use the refined poses to render and compare DRRs to real scans. Results can be found in Table S2 and Figure 4. *E.g.*, when zooming into the images of the latter figure, one can observe that DiffDRR suffers from pixelization, unlike GS methods.

NeurIPS Paper Checklist

1. Claims

Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

Answer: [Yes]

Justification: The abstract and introduction explicitly list our contributions.

Guidelines:

- The answer NA means that the abstract and introduction do not include the claims made in the paper.
- The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
- The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
- It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. Limitations

Question: Does the paper discuss the limitations of the work performed by the authors?

Answer: [Yes]

Justification: Please refer to Section 5 in the main paper for detailed discussion.

Guidelines:

- The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
- The authors are encouraged to create a separate "Limitations" section in their paper.
- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. Theory Assumptions and Proofs

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: Preliminaries are provided at the beginning of Section 3, and further theoretical background is provided in the supplementary material.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. Experimental Result Reproducibility

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: Implementation and experimental details are provided, so that an expert in the art should be able to reproduce the presented results.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general, releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
 - (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
 - (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
 - (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
 - (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. Open access to data and code

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [No]

Justification: All data used are public, and implementation release is pending approval at the moment.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. Experimental Setting/Details

Question: Does the paper specify all the training and test details (e.g., data splits, hyper-parameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: See Section 4.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. Experiment Statistical Significance

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [NA] .

Justification: See Section 4.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).

- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. Experiments Compute Resources

Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [Yes]

Justification: See implementation details in Section 4.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. Code Of Ethics

Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines>?

Answer: [Yes]

Justification: All standard guidelines were followed.

Guidelines:

- The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. Broader Impacts

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [Yes]

Justification: Please refer to Section 5 of the main paper for further details.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. Safeguards

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: [NA]

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. Licenses for existing assets

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification: Public datasets and libraries are referenced.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, paperswithcode.com/datasets has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset’s creators.

13. **New Assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [NA]

Justification: [NA]

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and Research with Human Subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Justification: [NA]

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification: [NA]

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.