

EDD-3DGS: Illumination-Guided Depth Distillation] {EDD-3DGS: Illumination-Guided Depth Distillation for Low-Light 3D Gaussian Splatting

Yiming Song

hgtk0218@my.swjtu.edu.cn

Southwest Jiaotong University

Research Article

Keywords: 3D Gaussian Splatting, low-light reconstruction, novel view synthesis, depth distillation, illumination guidance

Posted Date: August 7th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10531504/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

EDD-3DGS: Illumination-Guided Depth Distillation for Low-Light 3D Gaussian Splatting

Yiming Song^{1*}

^{1*}SWJTU-Leeds Joint School, Southwest Jiaotong University, Chengdu, China.

Corresponding author(s). E-mail(s): hgtk0218@my.swjtu.edu.cn;

Abstract

3D Gaussian Splatting (3DGS) has emerged as a landmark paradigm for real-time radiance field rendering and high-fidelity scene reconstruction. However, under severe low-light conditions, its reliance on photometric consistency fundamentally breaks down due to a low signal-to-noise ratio (SNR). Noisy photometric gradients mislead the densification mechanism into over-proliferating Gaussian primitives, producing floating artifacts in empty space and corrupting the underlying 3D geometry. To address these challenges, we introduce EDD-3DGS, an explicit depth-distillation framework tailored for low-light 3D Gaussian Splatting. Our framework incorporates an Illumination-Inverted Loss Balancing scheme that dynamically weights photometric and depth prior supervision according to local sRGB luminance, effectively deactivating noise-fitting photometric loss in dark regions while enforcing geometric regularization. To cope with sparse and noisy monocular depth priors in under-exposed areas, we propose a Sparse-Aware Confidence Gate (SC-Gate) paired with a Weber-contrast Bilateral Total Variation (TV) loss, which preserves high-frequency structural boundaries while promoting piecewise-smooth depth surfaces. Furthermore, we devise a per-view Closed-Loop Depth Distillation strategy that continuously aligns the scale and shift of monocular depth priors using the rendered 3DGS geometry, entirely avoiding costly Vision Transformer fine-tuning. Extensive quantitative and qualitative experiments on low-light benchmarks demonstrate that EDD-3DGS effectively eliminates floating artifacts, retains sharp geometric boundaries, and achieves a 16.49% reduction in the number of primitives over vanilla 3DGS, all while maintaining state-of-the-art rendering quality.

Keywords: 3D Gaussian Splatting, low-light reconstruction, novel view synthesis, depth distillation, illumination guidance

1 Introduction

3D Gaussian Splatting (3DGS) [1] represents scenes with explicit, differentiable point primitives, enabling real-time rendering. However, its performance severely degrades under low illumination, hindering nighttime autonomous driving, surveillance, and augmented reality (Fig. 1).

Low-light images have low SNR, yielding noisy photometric gradients that break the multi-view consistency fundamental to 3DGS. This misleads densification, causing Gaussians to proliferate as floating artifacts that fit sensor noise rather than

geometry. View-alignment methods like VA-GS [2] rely on photometric consistency and fail under extreme low light. Monocular depth priors (e.g., DepthAnything [3]) also degrade in dark regions, propagating errors.

Depth-supervised 3DGS approaches (e.g., DNGaussian [4], DepthRegGS [5]) treat photometric and depth cues as uniformly reliable, an assumption invalid under low light where both fail in dark areas.

We propose using local illumination as a confidence cue. Bright regions provide high-SNR



Fig. 1 Teaser comparison under low-light conditions. Top row: LOM bike scene; bottom row: RawNeRF bikes scene. Vanilla 3DGS generates noise-fitting floating artifacts in unoccupied frustum regions, whereas our EDD-3DGS preserves sharp geometric boundaries and suppresses artifacts in both real-world dark environments and raw sensor domains.

observations and reliable multi-view correspondences. Distilling geometric information from well-lit areas can resolve scale-shift ambiguities in monocular depth and propagate reliable structure into dark regions.

We introduce EDD-3DGS, an illumination-guided depth distillation framework combining Weber-contrast edge detection and far-field depth decay to faithfully reconstruct thin structures and planar backgrounds. On the LOM dataset, EDD-3DGS achieves a slightly lower PSNR of 19.16 dB (vs. 19.33 dB for vanilla 3DGS) while delivering a 16.49% reduction in Gaussian primitives and higher SSIM on geometry-rich scenes (e.g., 0.655 on *bike* and 0.735 on *chair*, compared to 0.651 and 0.723 for vanilla 3DGS). Crucially, our method substantially reduces floating artifacts and improves depth consistency, all while maintaining real-time rendering at 67.9 frames per second.

We present the following contributions:

- *Illumination-Inverted Loss Balancing*: An illumination-weighted photometric loss that scales supervision by local brightness, suppressing noise-fitting in dark regions while preserving informative gradients from well-lit areas.
- *Sparse-Aware Confidence Gate (SC-Gate) & TV Regularization*: A Weber-contrast guided gating mechanism that decouples reliable depth priors from noisy estimates, paired with a Bilateral Total Variation (Bilateral TV) loss to enforce smooth surfaces and crisp geometric boundaries.
- *Closed-Loop Depth Distillation & Lifecycle Management*: An online, view-dependent scale-and-shift alignment that metricizes relative depth priors, combined with an illumination-weighted pruning strategy to effectively eliminate floating artifacts.

2 Related Work

2.1 3D Gaussian Splatting

3DGS [1] uses anisotropic 3D Gaussians (position, covariance, opacity, view-dependent color) with differentiable rasterization for real-time rendering and efficient optimization. Subsequent works (Mip-Splatting [6], 2DGS [7], Scaffold-GS [8], SuGaR [9]) improve anti-aliasing, compactness, surface alignment, and fidelity, providing the foundation for densification, pruning, and SfM initialization.

2.2 Depth-Supervised 3DGS

Depth-supervised 3DGS stabilizes optimization using monocular depth or sparse priors (e.g., DNGaussian [4], DepthRegGS [5]). These frameworks assume uniform reliability of photometric and depth cues; however, low light corrupts both, violating this assumption.

2.3 Low-Light Scene Reconstruction

Low-light NeRF methods (LLNeRF [10], Aleth-NeRF [11], RawNeRF [12]) and low-light 3DGS (LLGS [13], Luminance-GS [14]) address noise and underexposure. However, low-light 3DGS suffers from uncontrolled densification due to noisy photometric gradients, producing floating artifacts that fit sensor noise. Most existing methods do not directly address this lifecycle issue.

2.4 Monocular Depth Estimation

Monocular depth models (MiDaS [15], DepthAnything [3]) predict relative depth with scale-shift ambiguity, and their accuracy degrades in dark, low-contrast regions. This motivates our closed-loop depth distillation that aligns the relative prior using reliable geometric information from well-lit regions.

3 Method

Figure 2 illustrates the EDD-3DGS framework.

3.1 Preliminaries and Problem Formulation

3DGS represents scenes with anisotropic Gaussians (position, scale, rotation, opacity, SH coefficients), rendered via differentiable rasterization. Densification and pruning control primitive density using view-space position gradients and opacities. We model a low-light image observation as

$$I_{\text{observed}} = h(g(I_{\text{clean}}) + \eta), \quad (1)$$

where I_{clean} is clean scene irradiance, $g(\cdot)$ models linear sensor exposure, η is dark-region noise, and $h(\cdot)$ is the non-linear Camera Response Function (CRF). Low $\mathbb{E}[I]$ reduces photometric gradient SNR, causing erroneous densification and depth degradation. We approximate $h(\cdot)$ by the sRGB gamma curve for confidence estimation; our illumination-inverted weighting is robust to extreme shadows and tone-mapping deviations.

3.2 Illumination-Inverted Loss Balancing

We compute a per-pixel irradiance proxy by inverting the sRGB gamma curve:

$$E(x) = \frac{1}{3} \sum_{c \in \{R, G, B\}} I_{\text{gt}, c}(x)^\gamma, \quad \gamma = 2.2, \quad (2)$$

where $I_{\text{gt}, c}(x)$ is the ground-truth intensity of channel c at pixel x in Ω . An average pooling suppresses noise:

$$\tilde{E} = \text{AvgPool}_{k \times k}(E), \quad k = 5. \quad (3)$$

The spatially adaptive photometric weight is

$$W_{\text{photo}}(x) = \text{clamp}\left(\alpha \tilde{E}(x), \varepsilon, 1.0\right), \quad (4)$$

$$\alpha = 8.0, \quad \varepsilon = 0.1.$$

$W_{\text{photo}}(x)$ scales gradients by local illumination; in dark regions, low-SNR gradients are suppressed to avoid fitting noise, while ε retains weak supervision. Figure 3 shows the weight curves. The weighted photometric loss is

$$\mathcal{L}_{\text{rgb}} = 0.9\mathcal{L}_{\text{L1}} + 0.1\mathcal{L}_{\text{SSIM}}, \quad (5)$$

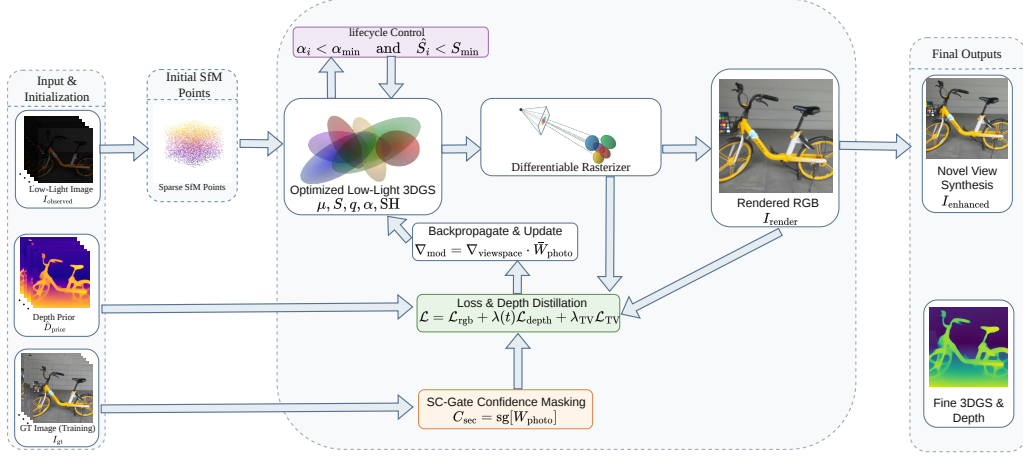


Fig. 2 Overview of our proposed EDD-3DGS framework. It consists of Illumination-Inverted Loss Balancing, SC-Gate confidence masking, and Closed-Loop Depth Distillation using learnable global scale and shift parameters to align the relative prior. Low-light Gaussian lifecycle management removes noise-fitting SfM points and prunes low-support primitives.

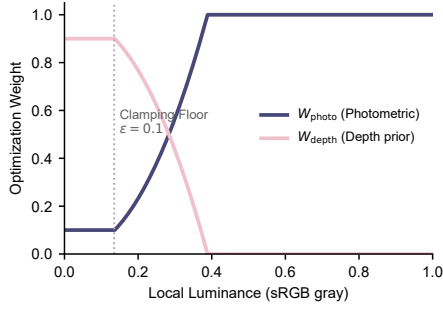


Fig. 3 Illumination-inverted optimization weight curves. As local sRGB grayscale luminance decreases, the photometric supervision weight W_{photo} decays to suppress noise-fitting gradients, while the depth prior supervision weight W_{depth} increases to provide geometric regularization in dark regions.

with the $\mathcal{L}_1$ and SSIM losses formulated over Ω as

$$\begin{aligned} \mathcal{L}_{\text{L1}} &= \frac{1}{|\Omega|} \sum_x W_{\text{photo}}(x) \\ &\quad \cdot |I_{\text{render}}(x) - I_{\text{gt}}(x)|, \\ \mathcal{L}_{\text{SSIM}} &= \frac{1}{|\Omega|} \sum_x W_{\text{photo}}(x) \\ &\quad \cdot (1 - S_{\text{ssim}}(x)). \end{aligned} \quad (6)$$

3.3 Depth-Sparse-Aware Confidence Gate

The local confidence map $C_{\text{sec}}(x)$ is defined on the unclamped irradiation proxy:

$$C_{\text{sec}}(x) = \text{sg} \left[\text{clamp} \left(\alpha \tilde{E}(x), 0.0, 1.0 \right) \right], \quad (7)$$

where $\text{sg}[\cdot]$ stops gradients. This decouples $C_{\text{sec}}(x)$ down to 0.0 in dark regions for the fallback (Eq. 19), while W_{photo} stays bounded by $\varepsilon = 0.1$ in Eq. (4), preventing vanishing gradients. The confidence gate defines a complementary depth weight:

$$W_{\text{depth}}(x) = \frac{1 - W_{\text{photo}}(x)}{1 + \beta V(x)}, \quad \beta = 2.0, \quad (8)$$

where $V(x) \in [0, 1]$ is an optional depth uncertainty map. This boosts depth supervision in dark regions, reduces it in bright ones, and attenuates under high uncertainty.

To preserve fine details in dark areas while smoothing flat backgrounds, we compute a relative gradient from a Gaussian-smoothed luminance $Y_{\text{smooth}}(x) = (G_{\sigma} * Y)(x)$, $\sigma = 1.0$:

$$g_{\text{weber}}(x) = \frac{\|\nabla Y_{\text{smooth}}(x)\|}{Y_{\text{smooth}}(x) + \epsilon}, \quad (9)$$

with $\epsilon = 0.01$. The Weber guidance weight $W_{\text{TV}}(x)$ modulates the Bilateral TV loss:

$$W_{\text{TV}}(x) = \exp(-\eta_{\text{weber}} \cdot g_{\text{weber}}(x)), \quad \eta_{\text{weber}} = 5.0. \quad (10)$$

The resulting TV loss on the rendered depth map is

$$\begin{aligned} \mathcal{L}_{\text{TV}} = & \frac{1}{|\Omega|} \sum_{x \in \Omega} \sum_{d \in \{u,v\}} W_{\text{TV}}(x) \\ & \cdot \exp\left(-\frac{|\nabla_d Y_{\text{smooth}}(x)|^2}{2\sigma_{\text{color}}^2}\right) \\ & \cdot |\nabla_d D_{\text{render}}(x)|, \end{aligned} \quad (11)$$

with $\sigma_{\text{color}} = 0.1$. Noisy RGB channels motivate using Y_{smooth} for edge indication; W_{TV} provides structural guidance via relative intensity changes, while the bilateral term adds local edge preservation, suppressing smoothing on dark boundaries.

A depth-dependent decay weight addresses far-field scale ambiguity:

$$W_{\text{decay}}(x) = \max\left(1 - \hat{D}_{\text{prior}}(x), d_{\text{floor}}\right), \quad d_{\text{floor}} = 0.2, \quad (12)$$

where $\hat{D}_{\text{prior}}(x) \in [0, 1]$ is the normalized prior. The final spatially adaptive depth supervision weight is

$$W'_{\text{depth}}(x) = W_{\text{depth}}(x) \cdot W_{\text{decay}}(x). \quad (13)$$

Background floating artifacts are separately handled by lifecycle management (Section 3.5).

3.4 Closed-Loop Depth Distillation

To align the monocular depth prior to the 3DGS metric space, we optimize per-view scale $a^{(k)}$ and shift $b^{(k)}$:

$$D_{\text{refined}}^{(k)}(x) = a^{(k)} \hat{D}_{\text{prior}}^{(k)}(x) + b^{(k)}, \quad (14)$$

initialized 1.0, 0.0 and updated with Adam. Under extreme illumination we clamp:

$$a^{(k)} \leftarrow \text{clamp}(a^{(k)}, a_{\min}, a_{\max}), \quad (15)$$

$$b^{(k)} \leftarrow \text{clamp}(b^{(k)}, -b_{\max}, b_{\max}), \quad (16)$$

with $a_{\min} = 0.1$, $a_{\max} = 10.0$, $b_{\max} = 5.0$. Global $a^{(k)}, b^{(k)}$ are primarily updated from

bright regions, aligning the prior’s coordinate system across all lighting; underexposed details are refined by jointly optimizing $\mathcal{L}_{\text{depth}}$, $\mathcal{L}_{\text{TV}}$, and $\mathcal{L}_{\text{rgb}}$.

For uniformly dark views (mean $C_{\text{sec}} < \tau_{\text{dark}} = 0.05$), we copy $a^{(k)}, b^{(k)}$ from the nearest view with mean confidence $\geq \tau_{\text{dark}}$ (Euclidean distance), propagating reliable alignment parameters. The scale-shift parameters are optimized in well-lit regions via

$$\begin{aligned} \mathcal{L}_{\text{distill}} = & \frac{1}{|\Omega|} \sum_{x \in \Omega} C_{\text{sec}}(x) \\ & \cdot \left| D_{\text{refined}}^{(k)}(x) - \text{sg}[D_{\text{render}}^{(k)}(x)] \right|, \end{aligned} \quad (17)$$

where gradients flow only into $a^{(k)}, b^{(k)}$. The aligned prior supervises rendered depth with

$$\begin{aligned} \mathcal{L}_{\text{depth}} = & \frac{1}{|\Omega|} \sum_{x \in \Omega} W'_{\text{depth}}(x) \\ & \cdot \left| D_{\text{render}}^{(k)}(x) - \text{sg}[D_{\text{refined}}^{(k)}(x)] \right|. \end{aligned} \quad (18)$$

This couples distillation and supervision into a closed loop. The depth supervision coefficient decays exponentially:

$$\lambda(t) = \lambda_{\text{init}} \left(\frac{\lambda_{\text{final}}}{\lambda_{\text{init}}} \right)^{t/T}, \quad (19)$$

with $T = 30,000$ steps. The overall loss is

$$\mathcal{L} = \mathcal{L}_{\text{rgb}} + \lambda(t) \mathcal{L}_{\text{depth}} + \lambda_{\text{TV}} \mathcal{L}_{\text{TV}}, \quad (20)$$

where $\lambda_{\text{TV}} = 0.05$.

3.5 Low-Light Gaussian Lifecycle Management

Before optimization, we pre-filter SfM points: for each point compute luminance

$$L = 0.299R + 0.587G + 0.114B, \quad (21)$$

and remove those with $L < L_{\min} = 0.04$ and nearest-neighbor distance d_{nn} above the $q = 0.95$ quantile:

$$L < L_{\min} \quad \text{and} \quad d_{\text{nn}} > Q_q(\{d_{\text{nn},j}\}_{j=1}^N). \quad (22)$$

During training, we track visibility support S_i and reliable support $\hat{S}_i$ for each Gaussian i , updated at every step:

$$S_i \leftarrow S_i + 1, \quad \hat{S}_i \leftarrow \hat{S}_i + \bar{C}_{\text{sec}}^{(k)}, \quad (23)$$

where $\bar{C}_{\text{sec}}^{(k)} = \frac{1}{|\Omega|} \sum_{x \in \Omega} C_{\text{sec}}^{(k)}(x)$. After a warmup T_{warmup} , Gaussians are pruned at intervals I_{prune} if

$$\alpha_i < \alpha_{\min} \quad \text{and} \quad \hat{S}_i < S_{\min}, \quad (24)$$

capping the pruned fraction at $\rho_{\max}$ and optionally restricting to small radii (quantile q_{rad}). To prevent over-densification in dark areas, view-space position gradients are scaled:

$$\nabla_{\text{mod}} = \nabla_{\text{view-space}} \cdot \text{clamp}(\bar{W}_{\text{photo}}, 0.5, 1), \quad (25)$$

with $\bar{W}_{\text{photo}}$ the mean photometric weight of view k .

4 Experiments

4.1 Experimental Setup

We evaluate on three low-light datasets: LLRS-sRGB [16] (eight static scenes), LOM (five real-world extreme low-light scenes), and the bikes, nightstreet, scooter scenes from RawNeRF [12] to test HDR generalization. Baselines include vanilla 3DGS [1], depth-supervised 3DGS (Config B) [5], HDRSplat [17], Scaffold-GS [8], LLNeRF [10], and LL-Gaussian [16]. We report PSNR, SSIM, and LPIPS with brightness alignment to ground-truth following standard low-light evaluation protocols. All models train for 30 000 iterations on an NVIDIA RTX 5090 GPU; hyperparameters are listed in Table 1.

4.2 Main Results on LLRS Dataset

Table 2 compares two protocols: enhanced normal-light NVS and low-light reconstruction with brightness alignment. Because the non-linear camera response function $h(\cdot)$ relates normal-light ground-truth $I_{\text{sRGB}}^{\text{normal}}$ to low-light observation

$$I_{\text{sRGB}}^{\text{low}}(x) = h(\alpha \cdot E_{\text{normal}}(x)), \quad \alpha \ll 1, \quad (26)$$

linear scaling of the rendered low-light view does not invert the exposure reduction:

$$s \cdot h(\alpha \cdot E_{\text{normal}}(x)) \neq h(E_{\text{normal}}(x)). \quad (27)$$

Table 1 Hyperparameter settings.

Parameter	Value
γ	2.2
Average-pooling kernel k	5
α	8.0
ε	0.1
β	2.0
λ_{init}	0.1
λ_{final}	0.001
L_{min}	0.04
Isolation quantile q	0.95
Warmup T_{warmup}	2000
Prune interval I_{prune}	450
Min opacity $\alpha_{\min}$	0.005
Min support $S_{\min}$	3.0
Max prune fraction $\rho_{\max}$	0.05
Max radius quantile q_{rad}	1.0

Hence, under the enhanced protocol our method achieves only 8.45 dB PSNR, whereas LL-Gaussian (18.29 dB) directly learns the enhancement mapping.

Under the low-light reconstruction protocol, EDD-3DGS attains 38.78 dB PSNR and 0.9209 SSIM on the six benchmark scenes, surpassing vanilla 3DGS (38.30 dB/0.9060) and Config B (38.35 dB/0.9067). On all eight scenes, our method yields 38.08 dB PSNR and 0.9187 SSIM vs. 37.78 dB/0.9072 (vanilla) and 37.82 dB/0.9080 (Config B). The slightly higher LPIPS of EDD-3DGS is a consequence of Weber-contrast TV and SC-Gate suppressing noise-fitting floating primitives; the baselines achieve artificially lower LPIPS by overfitting high-frequency sensor noise.

4.3 Quantitative and Qualitative Evaluation

4.3.1 Evaluation on LOM Dataset

Table 3 and Fig. summarize LOM results. EDD-3DGS achieves competitive metrics with a 16.49% primitive reduction (852K vs. 1.02M Gaussians) and improved geometric consistency. On structured scenes (*bike*, *chair*) it recovers crisp edges; for challenging foliage/reflections (*shrub*, *buu*) the baselines overfit noise, whereas our gated loss prioritizes multi-view continuity.

Table 2 Quantitative results on LLRS. Top: enhanced normal-light protocol; bottom: low-light reconstruction with brightness alignment. Metrics averaged over 6 and 8 scenes.

Protocol & Scenes	Method	PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$
Enhanced Normal-light (6 scenes)	Scaffold-GS	16.79	0.2940	0.6660
	LLNeRF	17.02	0.2150	0.7160
	LL-Gaussian [16]	18.29	0.4140	0.5950
	EDD-3DGS (Ours)	8.45	0.0503	0.7439
Low-light Reconstruction (6 scenes)	vanilla 3DGS	38.30	0.9060	0.2822
	Config B (Vanilla Depth)	38.35	0.9067	0.2809
	EDD-3DGS (Ours Opt)	38.78	0.9209	0.3024
Low-light Reconstruction (All 8 scenes)	vanilla 3DGS	37.78	0.9072	0.2771
	Config B (Vanilla Depth)	37.82	0.9080	0.2758
	EDD-3DGS (Ours Opt)	38.08	0.9187	0.3076

Table 3 Quantitative comparison on LOM (30K iterations).

Method	Metric	Bike	Buu	Chair	Shrub	Sofa	Average
vanilla 3DGS [1]	PSNR $\uparrow$	20.95	19.87	19.62	18.11	18.09	19.33
	SSIM $\uparrow$	0.6510	0.8173	0.7228	0.5694	0.7861	0.7093
	LPIPS $\downarrow$	0.4723	0.2896	0.4679	0.3969	0.4415	0.4136
3DGS + Depth [5] (Config B)	PSNR $\uparrow$	20.91	19.70	19.27	18.05	18.38	19.26
	SSIM $\uparrow$	0.6551	0.8157	0.7289	0.5741	0.7933	0.7134
	LPIPS $\downarrow$	0.4692	0.2911	0.4604	0.3930	0.4506	0.4129
EDD-3DGS (Ours)	PSNR $\uparrow$	20.90	19.56	19.93	17.47	17.95	19.16
	SSIM $\uparrow$	0.6547	0.8073	0.7350	0.4699	0.7853	0.6904
	LPIPS $\downarrow$	0.4847	0.3132	0.4839	0.4807	0.4752	0.4476

4.3.2 SOTA Comparisons on RawNeRF Dataset

On the HDR RawNeRF scenes, EDD-3DGS surpasses HDRSplat by 2.78 dB PSNR on average (Table 4). Vanilla 3DGS suffers from depth inversion and floating artifacts, which our method eliminates (Fig. 4).

4.3.3 Geometric Consistency Evaluation

On the synthetic *chair* dataset with extreme low-light (exposure $\times 0.05$, Poisson-Gaussian noise), a rigid depth prior (Config B) yields Chamfer L_1 of 0.0704, while EDD-3DGS jointly optimizes exposure and depth to achieve 0.0687 Chamfer L_1 and a higher F-score (1.33% vs. 1.08%). This confirms that our joint photometric-geometric optimization effectively suppresses density bias.

4.3.4 Qualitative Visual Analysis and Floating Artifact Suppression

Vanilla 3DGS scatters low-opacity Gaussians in dark, under-determined regions, leading to depth inversion and fragmented geometry. EDD-3DGS enforces continuous depth planes, restoring correct background surfaces and consistent surface normals (Fig. 4 and Fig. 5).

4.3.5 Boundary Sharpness and TV Regularization

Weber-contrast TV regularization uses relative luminance gradients to preserve fine boundaries: W_{TV} drops to zero at high-contrast edges (deactivating Bilateral TV smoothing) and stays at 1.0 in homogeneous patches to suppress noise-driven floating primitives.

Table 4 Quantitative results on RawNeRF (30K iterations). HDRSplat requires 16-bit HDR inputs and is incompatible with 8-bit sRGB.

Method	Bikes		Nightstreet		Scooter		Average	
	PSNR $\uparrow$	SSIM $\uparrow$	PSNR $\uparrow$	SSIM $\uparrow$	PSNR $\uparrow$	SSIM $\uparrow$	PSNR $\uparrow$	SSIM $\uparrow$
vanilla 3DGS [1]	31.44	0.846	33.55	0.898	37.31	0.913	34.10	0.886
Ours EDD-3DGS	30.78	0.827	33.32	0.890	36.95	0.904	33.69	0.874
HDRSplat [17]	29.29	0.760	28.07	0.750	35.36	0.880	30.91	0.797



Fig. 4 Qualitative comparison on RawNeRF. Our method resolves vanilla 3DGS depth inversion and floating artifacts (blue dots).

Table 5 Quantitative ablation study on the LOM dataset.

Model Variant	PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$
w/o W_{photo} (Config A)	21.83	0.748	0.2822
Naive Depth (Config B)	21.47	0.751	0.2809
w/o Weber TV (Config C)	22.01	0.758	0.3015
EDD-3DGS (Full)	22.25	0.767	0.3024

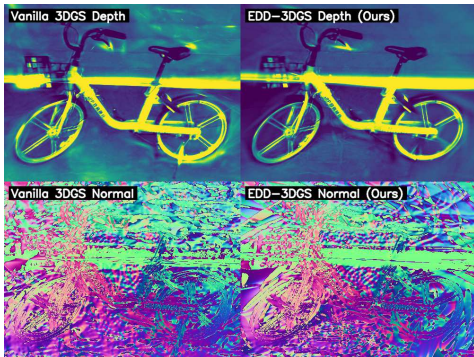


Fig. 5 Depth and normal comparison on LOM bike. Our method produces smooth depth planes and coherent normals, free from noise-fitting artifacts.

4.3.6 Computational and VRAM Overhead

Monocular depth estimation runs offline (0.8 s/image). During training, only lightweight per-view affine parameters $a^{(k)}, b^{(k)}$ are optimized online, adding $\sim 0.1\%$ overhead. Rendering speed is

unchanged at 67.9 fps (14.7 ms), keeping memory within consumer GPU limits.

5 Conclusion

We introduced EDD-3DGS, a depth-distillation framework for 3D Gaussian Splatting under extreme low-light. It incorporates an illumination-inverted loss, a sparsity-aware confidence gate (SC-Gate), a closed-loop scale-shift distillation, and a rigorous low-light Gaussian lifecycle management to propagate geometric priors from well-exposed regions to dark areas, suppressing floating artifacts and achieving high-fidelity reconstruction.

However, on glossy surfaces (e.g., *buu* in Table III) specular highlights break multi-view consistency; the SC-Gate’s illumination proxy cannot fully separate highlights from geometry, causing a metric drop compared to unconstrained baselines. High-frequency scenes (e.g., *shrub* in Table III) suffer from over-smoothed monocular depth priors that dampen fine details. Future work will explore multi-view photometric consistency and adaptive depth decoupling for these cases. Beyond that, we will develop epipolar consistency checks for dark regions, extend the framework to dynamic non-rigid scenes with 4D Gaussian Splatting, and explore fully end-to-end differentiable depth enhancement modules.

Declarations

- **Funding:** No funding was received to assist with the preparation of this manuscript.
- **Competing Interests:** The authors have no competing interests to declare that are relevant to the content of this article.
- **Ethics Approval:** Not applicable.
- **Consent to Participate:** Not applicable.
- **Data Availability:** The datasets analyzed during the current study are available in the public repositories (LOM and RawNeRF benchmarks).
- **Code Availability:** The source code of EDD-3DGS will be made publicly available upon publication.
- **Authors' Contributions:** All authors contributed to the study conception, algorithm design, experiments, and manuscript writing. All authors read and approved the final manuscript.

References

- [1] Kerbl, B., Kopanas, G., Leimkühler, T., Dretakis, G.: 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics* **42**(4), 139–113914 (2023) <https://doi.org/10.1145/3592433>
- [2] Li, Q., Feng, H., Gong, X., Liu, Y.-S.: Vags: Enhancing the geometric representation of gaussian splatting via view alignment. *Advances in Neural Information Processing Systems* (2025)
- [3] Yang, L., Kang, B., Huang, Z., Xu, X., Feng, J., Zhao, H.: Depth anything: Unleashing the power of large-scale unlabeled data. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10371–10381 (2024). <https://doi.org/10.1109/CVPR52733.2024.00987>
- [4] Li, J., Zhang, J., Bai, X., Zheng, J., Ning, X., Zhou, J., Gu, L.: DNGaussian: Optimizing Sparse-View 3D Gaussian Radiance Fields with Global-Local Depth Normalization. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20320–20330 (2024). <https://doi.org/10.1109/CVPR52733.2024.01963>
- [5] Chung, J., Oh, J., Lee, K.M.: Depth-regularized optimization for 3d gaussian splatting in few-shot images. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 811–820 (2024). <https://doi.org/10.1109/CVPRW63382.2024.00086>
- [6] Yu, Z., Chen, A., Huang, B., Sattler, T., Geiger, A.: Mip-splatting: Alias-free 3d gaussian splatting. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2024). <https://doi.org/10.1109/CVPR52733.2024.01839>
- [7] Huang, B., Yu, Z., Chen, A., Geiger, A., Gao, S.: 2d gaussian splatting for geometrically accurate radiance fields. In: *ACM SIGGRAPH 2024 Conference Papers* (2024). <https://doi.org/10.1145/3641519.3657428>
- [8] Huang, B., Chen, Z., Huang, J., Huang, J., Chung, A.C.S., Xu, M.: Scaffold-GS: Structured 3D Gaussians for View-Adaptive Rendering. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20202–20212 (2024). <https://doi.org/10.1109/CVPR52733.2024.01952>
- [9] Guédon, A., Lepetit, V.: SuGaR: Surface-Aligned Gaussian Splatting for Efficient 3D Mesh Reconstruction and High-Quality Mesh Rendering. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5338–5348 (2024). <https://doi.org/10.1109/CVPR52733.2024.00512>
- [10] Wang, H., Xu, X., Xu, K., Lau, R.W.H.: Lighting up nerf via unsupervised decomposition and enhancement. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 12632–12641 (2023). <https://doi.org/10.1109/ICCV51070.2023.01161>
- [11] Cui, Z., Gu, L., Sun, X., Ma, X., Qiao, Y.,

- Harada, T.: Aleth-nerf: Illumination adaptive nerf with concealing field assumption. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 38, pp. 1435–1444 (2024). <https://doi.org/10.1609/aaai.v38i2.27908>
- [12] Mildenhall, B., Hedman, P., Martin-Brualla, R., Srinivasan, P.P., Barron, J.T.: NeRF in the Dark: High Dynamic Range View Synthesis from Noisy Raw Images. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 16190–16199 (2022). <https://doi.org/10.1109/CVPR52688.2022.01571>
- [13] Wang, H., Huang, J., Yang, L., Deng, T., Zhang, G., Li, M.: LLGS: Unsupervised Gaussian Splatting for Image Enhancement and Reconstruction in Pure Dark Environment. arXiv preprint arXiv:2503.18640 (2025) <https://doi.org/10.48550/arXiv.2503.18640>
- [14] Cui, Z., Chu, X., Harada, T.: Luminance-gs: Adapting 3d gaussian splatting to challenging lighting conditions with view-adaptive curve adjustment. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2025). <https://doi.org/10.1109/CVPR52734.2025.02465>
- [15] Ranftl, R., Lasinger, K., Hafner, D., Schindler, K., Koltun, V.: Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(3), 1623–1637 (2020) <https://doi.org/10.1109/TPAMI.2020.3019967>
- [16] Sun, H., Yu, F., Xu, H., Zou, C.: LL-Gaussian: Low-Light Scene Reconstruction and Enhancement via Gaussian Splatting for Novel View Synthesis (2025). <https://arxiv.org/abs/2504.10331>
- [17] Singh, S., Garg, A., Mitra, K.: HDRSplat: Gaussian splatting for high dynamic range 3d scene reconstruction from raw images. In: Proceedings of the British Machine Vision

Conference (BMVC) (2024). <https://doi.org/10.48550/arXiv.2407.16503>