Novel View Synthesis Under Rainy Conditions with Neural Radiance Fields and Gaussian Splatting

Ivana Petrovska* and Boris Jutzi
Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany - ivana.petrovska@partner.kit.edu, boris.jutzi@kit.edu

Keywords: Neural Radiance Fields, Gaussian Splatting, Rain, Occlusions, Novel Views.

Abstract

Scene reconstruction and novel view synthesis from calibrated multi-view images still attracts a lot of attention in computer vision and graphics. However, the assumption that images are noise-free rarely holds in real-world scenarios where adverse weather conditions are inevitable. Being a part of our environment, we are particularly interested in rain as dynamic semi-transparent occlusion which imposes challenges to a complete and accurate geometry of the underlying features. More precisely, we qualitatively and quantitatively analyze the photometric image quality under rainy conditions generated by radiance field methods, namely: Neural Radiance Fields (NeRFs), 3D Gaussian Splatting (3DGS) and 2D Gaussian Splatting (2DGS) due to the different geometric representation. To assess the impact of rain to the scene reconstruction we consider raindrops and streaks captured with illumination variation as well as occlusion masks with different coverage. The evaluation is based on comparing 2D image metrics of the rendered novel views without and with masks. The experiments and results show that 3DGS achieves highest rendering fidelity in all scenarios without and with masks with SSIM of 0.724 and LPIPS of 0.291, followed by 2DGS with slightly lower scores, while NeRF exhibits lowest correspondence with the input images with SSIM of 0.584 and LPIPS of 0.384. We demonstrate the effectiveness of using masks to handle rain as transient element and radiance field methods’ ability to reliably approximate the geometry behind rain occlusions.

1. Introduction

Leveraging calibrated images for the task of 3D reconstruction and novel view synthesis still attracts a lot of attention in the computer vision community. With the development of Neural Radiance Fields (NeRFs) (Mildenhall et al., 2021) a scene is represented as a continuous volumetric field implicitly encoding occupied (and empty) space. NeRF intersects camera rays through the scene to generate a sampled set of 3D points and use those points and their corresponding 2D viewing direction as input to a neural network to produce an output set of density and radiance. The pixel color is calculated as a weighted average of radiance along the ray, with weights given by their accumulated volume density through gradient descent on the photometric loss between rendered and input views. However, the high computational overhead comes from sampling hundreds of points along the corresponding ray, querying density and then accumulating radiance through volume rendering. Despite accelerating training and inference up to 20x (Li et al., 2022, Li et al., 2023), densely rendering N query points along each camera ray is inefficient, since empty space that does not contribute to the rendered image is still sampled repeatedly with consequent noise.

Traditional graphics rendering paradigms don’t operate in this manner. Gaussian Splatting (GS) methods, conceived with 3D Gaussian Splatting (3DGS) (Kerbl et al., 2023) explicitly reconstruct scene geometry as a series of Gaussian ellipsoids initialized on a sparse point cloud, usually obtained as a by-product of Structure-from-Motion (SfM) (Schonberger and Frahm, 2016). Unlike NeRF-based methods which have a pre-determined and fixed computational complexity based on network hyperparameters, the size of learnable parameters is dynamically adjusted to align with the complexity of the underlying 3D scene. More specifically, 3DGS uses an adaptive density control strategy to flexibly increase or decrease the number of Gaussians for optimizing reconstruction, leading to variable computational complexity in terms of GPU memory occupancy and training time cost. During training, the gradient-based optimization fits the spherical harmonic (SH) coefficients, describing the color, to match real-world appearance. After depth-sorting of the Gaussians, the volumetric alpha blending integrates alpha-weighted appearance from front to back until the accumulated opacity reaches saturation. One of the input views is selected each iteration and a novel view is rendered w.r.t. the camera poses. Then, a photometric loss is computed between the rendered view and the input image. Shifting from volumetric ray marching and tracing to differentiable forward-mapping rasterization omitting computing points in empty space, 3DGS significantly reduces computational overhead achieving real-time radiance field rendering while maintaining visual quality competitive with NeRF methods.

Nevertheless, 3DGS evaluates a Gaussian’s value at the intersection between a pixel ray and a 3D Gaussian, leading to inconsistent depth when rendered from different viewpoints. To better represent local surface patches, 2D Gaussian Splatting (2DGS) (Huang et al., 2024) explicit ray-splat intersection results in better reconstruction quality. With depth regularization the 2D Gaussian surfels in different views are constrained to have consistent depth, preventing floaters. Furthermore, inferring surface normals from depth maps enforces normal consistency across views preventing discontinuities to ensure that 2D Gaussians from different viewpoints obey the same surface orientation. Embedded in 3D space, the view-dependent surfels are optimized in screen space per frame.

However, the assumption that images are perfectly posed and noise-free rarely holds in real-world scenarios where adverse

* Corresponding author

weather conditions are inevitable. Being a part of our environment, we are particularly interested in rain as semi-transparent dynamic occlusion which degrades captured image quality, thus it can be challenging to generate a complete and accurate geometry of the underlying features.

In this contribution we qualitatively and quantitatively analyze the photometric image quality under rainy conditions generated by radiance field methods, namely: NeRFs, 3DGS and 2DGS due to the different geometric representation. Our investigations are based on real-world scenarios under rainy conditions captured with camera illumination variations, tackling raindrops and streaks to investigate how different type and level of rainfall affect the geometric reconstruction. To assess the impact of rain to the scene representation we restrict the rain by considering occlusion masks with different coverage. To compare how the methods perform without visibility obstructions, we include an indoor occlusion-free scenario. The evaluation is based on comparing 2D image metrics of the rendered novel views without and with masks.

Our contributions are three-fold:

- We provide a qualitative and quantitative analysis of the photometric image quality among NeRF, 3DGS and 2DGS under rain conditions, considering illumination variations.
- We show that radiance field methods can reliably approximate the geometry behind rain occlusions.
- We demonstrate that 3DGS outperforms in novel view synthesis without and with masks showing robustness under rainy conditions.

The contribution is organized as follows. In Section 2 an overview of radiance field reconstruction under rainy conditions is presented. In Section 3 the applied 3D reconstruction methods and the evaluation metrics are given. The used dataset along with implementation details are provided in Section 4. The qualitative and quantitative results are presented in Section 5. The discussion is laid out in Section 6 and Section 7 gives the concluding insights.

2. Related Work

In the following, we summarize radiance field geometric reconstruction tackling occlusions, specifically under rainy conditions for NeRFs (Section 2.1) and GS (Section 2.2) respectively.

2.1 Neural Radiance Fields

Given multi-view posed images, NeRFs (Mildenhall et al., 2021) use stochastic gradient descent to optimize density and color by minimizing photometric discrepancy between rendered and ground truth images. Addressing scene reconstruction through occlusions, RobustNeRF (Sabour et al., 2023) and Occlusion-Free NeRF (Zhu et al., 2023) offer robust optimization with distractors leading to an occlusion-free scene. With the prior that occluded regions are at least partially visible from other viewpoints, OccludeNeRF (Shi et al., 2025) and DeclutterNeRF (Shi et al., 2025) leverage NeRFs multi-view consistency, joint camera optimization, occlusion annealing regularization and stochastic structural similarity loss with information sharing among the views. Tackling reconstruction in adverse weather conditions, DehazeNeRF (Chen et al., 2024b) and ScatterNeRF (Ramazzina et al., 2023) decompose the scene into clear and participating media for real-world scenarios. DerainNeRF (Li et al., 2024) addresses the challenge of masking raindrops by leveraging a NeRF-based network and a pre-trained detector leading to rain-free reconstruction. RainyScape (Lyu et al., 2024) extends the reconstruction to rain streaks.

2.2 Gaussian Splatting

Starting from sparse points usually produced during camera calibration, the scene is represented with 3D Gaussians containing position, rotation, scale, opacity and color (Kerbl et al., 2023). The interleaved optimization/density control of the Gaussians, notably optimizing anisotropic covariance aims to achieve an accurate representation of the geometry. Extending GS to uncontrolled in-the-wild settings containing occluders, WildGaussians (Kulhanek et al., 2024), Splatfacto-W (Xu et al., 2024) and SpotlessSplats (Sabour et al., 2024) exploit semantic features to effectively identify distractors without explicit supervision. However, they under-perform in challenging scenarios where occlusions are present in all training images. OccluGaussian (Liu et al., 2025) clusters training cameras based on their positions and co-visibilities to acquire multiple regions. A region-based rendering technique accelerates large scene rendering by culling Gaussians invisible to the region where the viewpoint is located. Addressing the complexities of 3D reconstruction under rainy conditions in synthetic and real-world setup, WeatherGS (Qian et al., 2024) and DeRainGS (Liu et al., 2024) learn rain-free novel views from rainy input images with raindrops and streaks. Enabling a comprehensive evaluation of reconstruction quality with different camera focus, Raindrop GS (Teng et al., 2025) employs 3DGS variants using raindrop-focused, background-focused and rain-free input images.

We can conclude that there is a significant quantity of NeRF and GS methods focusing on rain-free reconstruction and generating novel views. However, in all of the above-mentioned methods the rain is weak and the droplets are small with regular circular shape and the rain occupies small image parts. In contrast, our two rainy scenarios depict heavier rainfall with different type of rain and challenging irregular shape in form of elongated streaks. In addition, the images are captured once without and once with camera flash from the same position to allow investigating of the impact of different camera illumination variations to the geometric reconstruction and subsequently the novel view synthesis.

3. Methodology

As depicted in Figure 1, we describe the applied 3D radiance field reconstruction methods in Section 3.1. The 2D evaluation metrics are explained in Section 3.2.

3.1 3D Reconstruction Methods

Neural Radiance Fields (NeRFs) NeRF (Mildenhall et al., 2021) employs a multilayer perceptron (MLP) neural network F_Θ to parameterize a 5D input (3D position $\mathbf{x} = (x, y, z)$ and 2D viewing direction $\mathbf{d} = (\theta, \phi)$) into density σ and color $\mathbf{c} = (r, g, b)$.

$$F_\Theta : (\mathbf{x}, \mathbf{d}) \rightarrow (\sigma, \mathbf{c}) \tag{1}$$

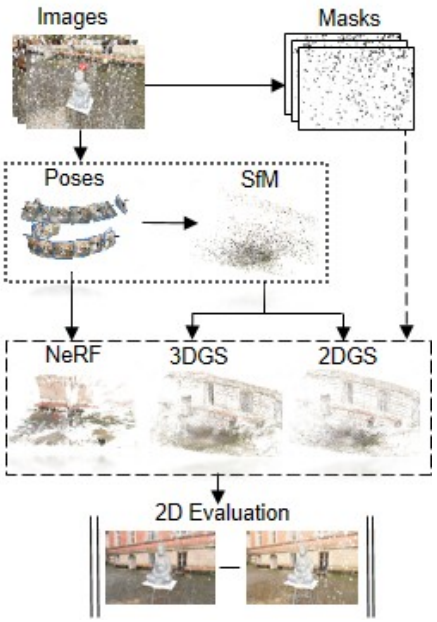

Figure 1. Using rainy images as input, the camera poses along with the sparse point cloud are estimated through SfM. The images along with the poses are input to NeRF, while 3DGS and 2DGS additionally require the sparse point cloud for initialization (Section 3.1). By considering occlusion masks we restrict the rain from the optimization and evaluate the novel views without and with masks (Section 3.2).

A volume rendering integrates a set of functions of σ and c over a ray, constructed from the projection center through the scene first with equidistant sampling and second, in more detailed manner, according to a probability density function, centered at the highest density response from the former sampling. The volume rendering yields a color representation at this particular position in the rendered image.

For rendering novel views, NeRF casts a ray $\mathbf{r}(t) = \mathbf{o} + t\mathbf{d}$ for each pixel, where $\mathbf{o}$ is the camera origin and $\mathbf{d}$ is the ray direction. The expected color $C(\mathbf{r})$ of the ray is computed by integrating the radiance along the ray.

$$C(\mathbf{r}) = \int_{t_n}^{t_f} T(t)\,\sigma(\mathbf{r}(t))\,\mathbf{c}(\mathbf{r}(t), \mathbf{d})\,dt \tag{2}$$

where $T(t) = \exp\Big(-\int_{t_n}^{t} \sigma(\mathbf{r}(s))\,ds\Big)$ is the accumulated transmittance and $[t_n, t_f]$ are the near and far bounds of the ray. The continuous integral is approximated using numerical quadrature.

$$\hat{C}(\mathbf{r}) = \sum_{i=1}^{N} T_i\,(1 - \exp(-\sigma_i\delta_i))\,\mathbf{c}_i \tag{3}$$

where $T_i = \exp\Big(-\sum_{j=1}^{i-1} \sigma_j\delta_j\Big)$, $\delta_i = t_{i+1} - t_i$ is the distance between adjacent samples and N is the number of samples along the ray.

After the execution of the $\mathcal{L}$ loss function that takes the rendered and ground truth image, a backward pass is executed. Whenever the network predicts non-zero density and a different color from the background, the gradients will encourage the network to update its predicted radiance. With the occlusion binary masks $M(u, v) \in [0, 1]$, we train NeRF by restricting the masked pixels not to contribute to the optimization in the photometric loss.

3D Gaussian Splatting (3DGS) Given a set of 3D Gaussian primitives (Kerbl et al., 2023) parametrized by mean $\mu \in \mathbb{R}^3$, anisotropic covariance $\Sigma \in \mathbb{R}^{3\times3}$, color c via SH and opacity α, first the projected mean $\mu' \in \mathbb{R}^2$ and covariance $\Sigma' \in \mathbb{R}^{2\times2}$ on the image planes are calculated. Then, the Gaussians are sorted such that all within the bounds of a tile are grouped and ordered by depth z relative to the camera center and each pixel within the tile is rendered with alpha-compositing.

The 3D covariances are reparametrized with a scaling matrix $S = \text{diag}(s) \in \mathbb{R}^{3\times3}$ represented by a scale vector $s \in \mathbb{R}^3$ and a rotation matrix $R \in \mathbb{R}^{3\times3}$ represented by a rotation quaternion $q \in \mathbb{R}^4$.

$$\Sigma = RSS^TR^T$$

The projection of 3D Gaussians is approximated with the Jacobian of the perspective projection equation.

$$J = \begin{bmatrix} \frac{f_x}{z} & 0 & -\frac{f_x x}{z^2} \\ 0 & \frac{f_y}{z} & -\frac{f_y y}{z^2} \\ 0 & 0 & 0 \end{bmatrix}$$

$$\Sigma' = JW\Sigma W^TJ^T$$

where $[W|t]$ is the world-to-camera transformation matrix and f_x, f_y denotes the camera focal lengths.

The color c of each pixel is computed by blending contributions from N overlapping Gaussians.

$$C = \sum_{j\in N} c_j\alpha_jT_j \tag{4}$$

where

$$T_j = \prod_{k=1}^{j-1}(1 - \alpha_k)$$

represents the accumulated transparency. Here, c_j and α_j denote the color and opacity of the j-th Gaussian respectively (Ye et al., 2025).

The 3D Gaussians are trained to minimize the photometric loss between rendered and input views. With the occlusion masks the gradients are restricted to only learn from the unmasked pixels.

2D Gaussian Splatting (2DGS) Each Gaussian surfel (Huang et al., 2024) is parameterized with two principal tangential axis vectors $\mathbf{t}_u$ and $\mathbf{t}_v$ paired with two scaling factors s_u and s_v, the surfel center $\mathbf{p}$, opacity α and the SH parameters $\mathbf{c}$. The surfel normal $\mathbf{n}$ is defined by $\pm(\mathbf{t}_u \times \mathbf{t}_v)$, pointing towards the viewing direction.

Given a pixel $\mathbf{x}$ in image space, 2DGS computes precise ray-surfel intersection (u, v) by the projection matrix $\mathbf{P}$ and the splat-to-world transformation $\mathbf{H}$.

$$\mathbf{x} = (xz, yz, z, z)^\top = \mathbf{PH}(u, v, 1, 1)^\top \tag{5}$$

the transformation $\mathbf{H}$ is defined by the surfel parameters.

$$\mathbf{H} = \begin{bmatrix} s_u\mathbf{t}_u & s_v\mathbf{t}_v & \mathbf{0} & \mathbf{p} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6}$$

Then the surfel weight is computed by querying a normalized Gaussian distribution.

$$G(\mathbf{x}) = \exp\left(-\frac{u^2 + v^2}{2}\right)$$

The pixel color $\mathbf{c}$, depth D and surface normal $\mathbf{N}$ are then accumulated across all intersected surfels along the ray from front to back by alpha-blending.

$$\mathbf{C}(\mathbf{x}) = \sum_i c_i(d)\alpha_i G_i(\mathbf{x}) T_i \tag{7}$$

$$D(\mathbf{x}) = \sum_i z_i(\mathbf{x})\alpha_i G_i(\mathbf{x}) T_i \tag{8}$$

$$\mathbf{N}(\mathbf{x}) = \sum_i \mathbf{n}_i\alpha_i G_i(\mathbf{x}) T_i \tag{9}$$

where $T_i = \prod_{j=1}^{i-1}(1 - \alpha_j G_j)$ is the transmittance, and $\mathbf{d}$ is the viewing direction. The photometric loss is enforced with depth, to align splats with surface geometry and normal regularization to ensure consistent orientation across views.

$$L_d = \sum_{i,j} \alpha_i G_i T_i \alpha_j G_j T_j (z_i - z_j)^2 \tag{10}$$

$$L_n = \sum_i \alpha_i G_i T_i (1 - \mathbf{n}_i \mathbf{N}_d) \tag{11}$$

Subsequently, considering the occlusion masks the algorithm is guided not to propagate for the masked image pixels.

3.2 2D Evaluation Metrics

Peak Signal-to-Noise Ratio (PSNR) It is defined in terms of the logarithmic ratio between the maximum possible signal intensity and the Mean Squared Error (MSE) between reference and test images. Measured in decibel (dB), higher value indicates better reconstruction quality with less noise relative to the reference views (Wang and Bovik, 2009).

Based on the MSE between a reference image I and a test image K of size $m \times n$

$$\mathrm{MSE}(I, K) = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \big(I(i,j) - K(i,j)\big)^2 \tag{12}$$

PSNR (Hore and Ziou, 2010) is defined as

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{L^2}{\mathrm{MSE}}\right) = 20 \cdot \log_{10}\left(\frac{L}{\sqrt{\mathrm{MSE}}}\right) \tag{13}$$

where L is the maximum possible pixel intensity (e.g. 255 for 8-bit images).

Structural Similarity Index Measure (SSIM) Unlike pixel-wise metrics such as PSNR, SSIM (Wang et al., 2004) is designed to closely align with human visual perception by focusing on structural consistency. SSIM measures the perceptual similarity between two images comparing structural information, luminance and contrast. Its values range from -1 to 1, where higher values indicate better similarity to the reference image.

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)\,(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)\,(\sigma_x^2 + \sigma_y^2 + C_2)}, \tag{14}$$

where μ_x, μ_y are local means, σ_x^2, σ_y^2 variances, σ_{xy} covariance, and C_1, C_2 stabilizing constants.

Learned Perceptual Image Patch Similarity (LPIPS) In contrast to PSNR and SSIM, which rely on pixel intensity or structural information, LPIPS (Zhang et al., 2018) aligns more closely with human perception to evaluate image similarity by comparing features extracted from pretrained neural networks. Since it quantifies the distance between feature maps between images, lower value indicates perceptually closer alignment.

LPIPS is a deep learning based perceptual metric that compares features extracted from pretrained neural networks.

$$d(x, y) = \sum_l \frac{1}{H_l W_l} \sum_{h,w} \Big\| w_l \odot \big(\hat{F}_l(x)_{h,w} - \hat{F}_l(y)_{h,w}\big) \Big\|^2 \tag{15}$$

where $F_l(\cdot)$ denotes the feature map at layer l, $\hat{F}_l$ are unit-normalized features, w_l are learned per-channel weights, and (h, w) are spatial coordinates over a feature map of resolution $H_l \times W_l$.

4. Experiments

We first provide details about the used dataset (Section 4.1), followed by the implementation details for each 3D reconstruction method (Section 4.2) and evaluation metrics (Section 4.3).

4.1 Dataset

Our investigations are based on two real-world outdoor scenarios from the STELLA dataset (Petrovska and Jutzi, 2024, Petrovska and Jutzi, 2025), namely *General-Rain* where the images are captured without camera flash and *Illuminated-Rain* containing corresponding images captured with flash depicting rainfall in front of a Buddha statue. For comparability we also include the indoor occlusion-free scenario *Original*. Each scenario consists of 125 images with 1840x1228 px resolution. The binary masks (24% masked pixels in *General-Rain* and

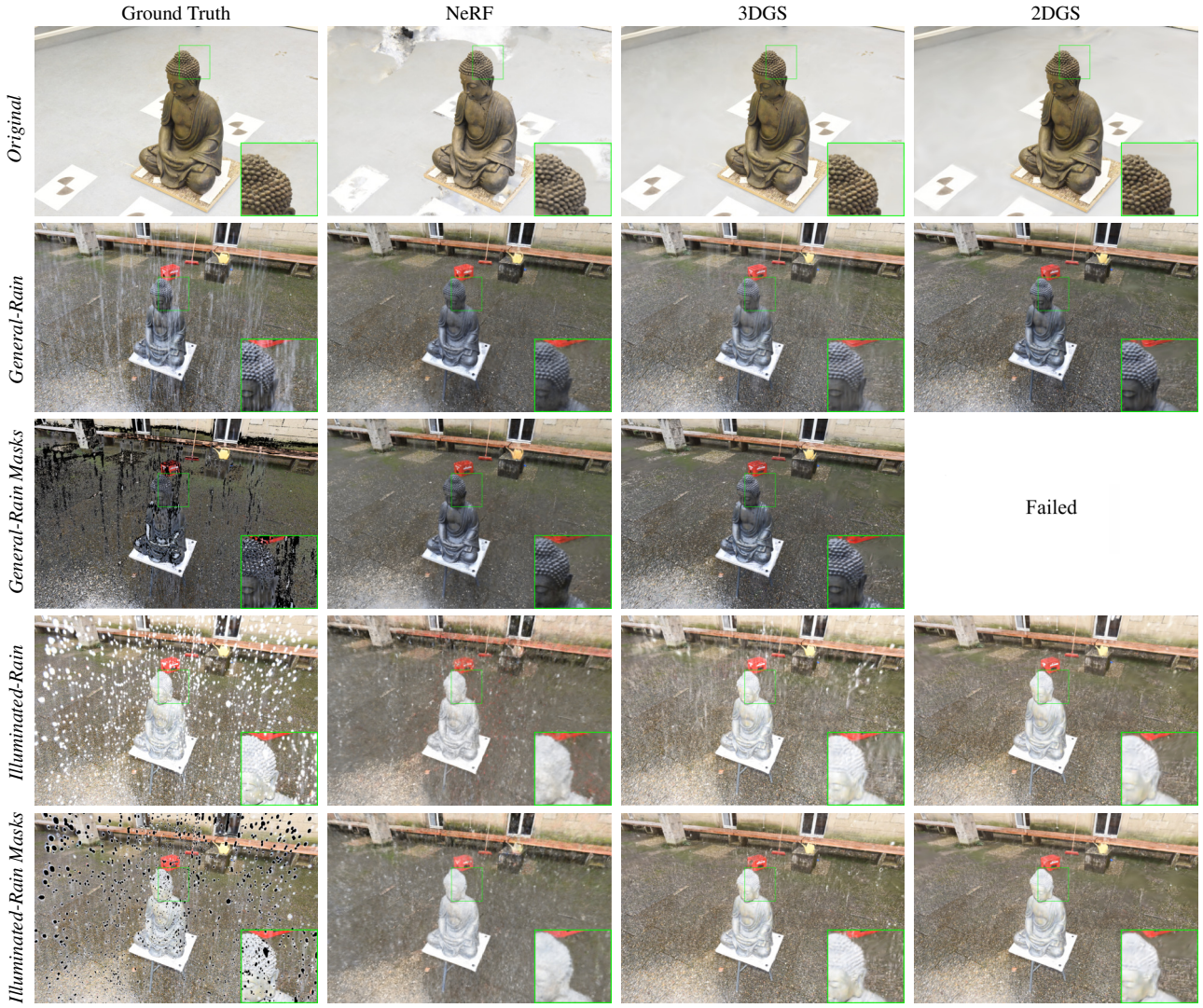

Figure 2. Novel view synthesis. The first column displays the ground truth (black pixels correspond to masked parts), while the second, third and fourth columns showcase the rendering results for each method. 3DGS achieves high rendering fidelity under rainy conditions, while NeRF shows lowest visual correspondence with the input images. Masking the rain leads to cleaner scene reconstruction and renderings.

4% in *Illuminated-Rain*) which indicate areas that shouldn't be sampled are in the same resolution and .png format as the input images. The camera poses along with the sparse point cloud are estimated in COLMAP (Schönberger et al., 2016).

4.2 Implementation Details

NeRF As NeRF representative method we train Nerfacto in Nerfstudio (Tancik et al., 2023) v1.1.0 with the default parameters. The neural network with 64 neurons per layer, samples 256 uniformly distributed positions on a ray performed by a piecewise sampling. From those samples, a first refinement iteration is applied using the proposal sampling with 96 samples followed by a second iteration with 48 samples. The pose refinement is disabled for a fair comparison.

3DGS & 2DGS We use the default parameters with learning rates of 0.0025 for SH, 0.05 for opacity adjustments, 0.005 for scaling operations, 0.001 for rotation and 0.0002 threshold for positional gradient for Gaussian densification which starts from the 500th until 15.000 iteration, every 100 iterations. The opacity is set close to zero every 3.000 iterations to prevent floaters close to the cameras. The SfM point cloud for each scenario is used for initialization. 3DGS and 2DGS as well as NeRF are trained 30.000 iterations and every 8th frame is taken for test, on a desktop PC with an Intel i9 CPU with 32GB RAM and Nvidia RTX3090 GPU.

4.3 Evaluation

For a comprehensive evaluation, besides the radiometry metrics we provide geometric quantification of the reconstruction by calculating Mean Error (Mean), Standard Deviation (SD) and Root Mean Squared Error (RMSE) and Completeness with 5mm distance threshold w.r.t a reference mesh from Structured Light of the Buddha statue.

5. Results

We present the results qualitatively through the rendered novel views (Figure 2) and quantitatively through photometric evaluation (Table 1) and geometric accuracy and completeness (Table 2) for each method.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-1-2026-17-2026 | © Author(s) 2026. CC BY 4.0 License.

Table 1. 2D image metrics of rendered novel views among NeRF, 3DGS and 2DGS across the rainy scenarios, without and with masks. GS methods significantly outperform NeRF, with 3DGS achieving best rendering fidelity. The first, second and third best results are highlighted.

Scenario	Method	PSNR ↑	SSIM ↑	LPIPS ↓
Original	NeRF	24.38	0.902	0.165
	3DGS	29.52	0.946	0.185
	2DGS	28.17	0.938	0.204
General-Rain	NeRF	17.86	0.489	0.490
	3DGS	21.15	0.624	0.352
	2DGS	19.60	0.594	0.390
General-Rain Masks	NeRF	18.77	0.802	0.165
	3DGS	19.56	0.839	0.165
	2DGS	Failed		
Illuminated-Rain	NeRF	15.90	0.314	0.692
	3DGS	18.86	0.592	0.417
	2DGS	18.04	0.653	0.446
Illuminated-Rain Masks	NeRF	16.12	0.414	0.408
	3DGS	18.63	0.620	0.338
	2DGS	18.77	0.612	0.350
Average	NeRF	18.61	0.584	0.384
	3DGS	21.54	0.724	0.291
	2DGS	21.14	0.677	0.348

Generally, NeRF shows the lowest rendering quality compared to the GS methods in all scenarios, even in *Original* which is occlusion-free with PSNR of 24.38dB and SSIM of 0.902. This is mostly due to the background and floor which are blurry as during image capturing the camera focus is always on the object. Logically, the scores decrease with occlusions, namely in *General-Rain* to 17.86dB and 0.489 accordingly. The rain streaks appear smeared and less distinct due to blurring and partial absorption into the background, also with masks. In *Illuminated-Rain* due to the higher camera exposure, the rendering quality is worse and lowest among the methods with PSNR of around 15dB and structural similarity slightly above 0.3. The masks aid in removing some of the raindrops, however bright blobs and slightly erroneous highlights are still present in the scene. In rainy conditions NeRF tends to blur or attenuate the rain streaks and partially blends them into background, indicating the inherent difficulty in cleanly separating transient components. Inline with this, the 3D geometric accuracy is the lowest among the methods by a large margin with RMSE of 13.67mm, but it shows robustness for a complete reconstruction with highest point coverage of 66.39%.

3DGS achieves the best rendering fidelity. In *Original* the novel views match the input images with highest PSNR of 29.52dB and SSIM of 0.946. Challenged by the semi-transparent and dynamic rain occlusions, the scores decrease but are still highest among the methods. In both rainy scenarios without masks, the rain streaks are partially reconstructed meaning better depth separation between rain and the rest of the scene, unlike NeRF and 2DGS which instead create blurry artifacts. The binary masks lead to rain-free novel views, although parts of the floor are blurred. The highest average rendering scores with LPIPS of 0.291 is complemented by second best 3D accuracy reaching RMSE of 8.51mm, however lowest completeness percentage of 43.06. 3DGS shows a good balance between geometry preservation and rendering of semi-transparent particles improving compositing of rain against background geometry.

2DGS shows competitive rendering quality in *Original* with PSNR of 28.17dB and SSIM of 0.938. Nevertheless, it’s unable to reconstruct the rain in *General-Rain* as due to its dynamic nature it’s difficult to find precise ray-splat intersection. With masks it runs Out Of Memory (OOM) caused by an explosion in Gaussian count (Petrovska and Jutzi, 2025) challenged by the complex light transmission properties of the rain as it assumes surfaces with full opacity. In *Illuminated-Rain* and *Illuminated-Rain Masks* the results are somewhat similar and moderate just behind 3DGS, since raindrops don’t fully transmit light the regions behind rain exhibit localized intensity attenuation which produces visible artifacts. However, 2DGS shows lowest geometric errors with average RMSE of 7.92mm and moderate completeness of 55.42%.

Table 2. Average 3D accuracy and completeness. 2DGS reaches highest accuracy, while NeRF although shows poor correspondence with the ground truth, maintains stable point coverage. The first, second and third best results are highlighted.

	Method	Accuracy [mm]			Completeness
		Mean ↓	SD ↓	RMSE ↓	[%] ↑
Average	NeRF	-7.27	11.56	13.67	66.39
	3DGS	-2.28	8.15	8.51	43.06
	2DGS	-1.56	7.76	7.92	55.42

In summary, GS methods show better geometric reconstruction and rendering quality than NeRF in all scenarios without and with masks. 3DGS achieves highest photometric and second best geometric accuracy. 2DGS rendering fidelity is slightly lower, however it shows closest correspondence with the ground truth mesh in 3D. NeRF produces plausible appearance, but smooths out the rain, resulting in floaters and splotchy artifacts that degrade the reconstruction, remaining highest point coverage. The rain is not reconstructed in the scene (Figure 3) and subsequently not rendered in the novel views due to its dynamic nature by all methods. The raindrops and streaks are not fixed in 3D space, thus appear at different positions in each image. Moreover, the rain acts as semi-transparent dynamic reflections that can blend into the background and may be misinterpreted as actual scene geometry, leading to incorrect depth cues. The camera illumination during image acquisition also affects the scene representation. Without flash the rain appears like a curtain and has a foggy effect. With flash, the rain is more pronounced and appears as bright, high-contrast spots creating distinct occlusions imposing bigger challenges for the methods. Applying the masks leads to visually cleaner reconstruction without rain indicating improved scene consistency. The fact that the image metrics, especially SSIM and LPIPS improve with masks suggests that the rendered scene is perceptually cleaner. The rendering quality is also affected by the binary masks which have limitations; the elongated rainfall effects are not masked, especially in *General-Rain* due to the low contrast and semi-transparency; object parts, background and floor are masked and some rain streaks are missed (Petrovska and Jutzi, 2025), creating inconsistency across views.

6. Discussion

NeRF and GS methods compare rendered with input images and the optimization is based on comparing color values of the pixels, trying to match the input views. Representing geometry as a continuous volumetric density field, NeRF assumes static scenes expecting consistent appearance across views. The semi-transparency of the rain breaks this assumption, so the inconsistent pixel color is reconstructed with blurry geometry and

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-1-2026-17-2026 | © Author(s) 2026. CC BY 4.0 License.

Figure 3. Point cloud reconstruction for both rainy scenarios. Due to its dynamic nature the rain is not reconstructed in the scene, but impacts the geometry of the features it occludes.

floaters (Figure 2). The algorithm may incorrectly assign rain artifacts to 3D structure, leading to foggy floating geometry. Thus, the rain in both scenarios in the rendered images is blurry and the novel view synthesis quality is lowest in rainy conditions. On the other hand, GS methods achieve higher correspondence with the input images. 3DGS shows better rendering fidelity than 2DGS, nonetheless the latter exhibits better 3D geometric accuracy (Table 2). We argue that this trade-off is due to the geometric regularization constraints compelling Gaussian kernels to distribute along object surfaces, which benefits geometric reconstruction but compromises rendering quality (Yang et al., 2025, Chen et al., 2024a). The large difference in PSNR, SSIM and LPIPS between GS and NeRF is particularly notable, suggesting that the structural integrity and perceptual quality of GS reconstruction are vastly superior, especially under rainy conditions. This implies that GS methods are more robust to artifacts introduced by rain and complex lighting.

It is worth noting that due to the lack of rain-free images, the photometric quantitative evaluation is limited to consistency in non-occluded regions: is the quality of areas that aren’t affected by rain preserved when using masks? Same as when training, we apply the masks on the input images and don’t take the masked black pixels for 2D metric calculation. Therefore, the LPIPS for NeRF and 3DGS in *General-Rain Masks* is 0.165 (Table 1) and better than the occlusion-free *Original* scenario because the averaging or blurring effects that can occur when reconstructing all regions uniformly is avoided. LPIPS is computed only on visible (non-masked) regions, on a curated subset of regions, while *Original* includes all pixels, potentially including more challenging regions that lower the overall score.

For all methods the performance drops in *Illuminated-Rain* as the rain captured with flash appears as bright, high-contrast spots creating distinct occlusions which are more complex to model than the rain streaks in *General-Rain* without flash, leading to lower photometric image quality. With masks, the reconstruction is better suggesting that separating the rain/illumination artifacts is effective in complex rainy conditions. Moreover, the scene is reconstructed with 11 images less which are not taken during pose estimation phase due to homogeneous background and texture repetition additionally challenged by the rain occlusions (Petrovska and Jutzi, 2025). In both rainy scenarios, the rain is not reconstructed in the SfM point cloud because it’s dynamic and inconsistent across views (Figure 3). Although radiance field methods optimize on images on which the rain is visible can’t find matches because it’s different in each image.

7. Conclusion

In this contribution, we provide a qualitative and quantitative analysis of the novel view synthesis under rainy conditions reconstructed by radiance field methods: NeRF, 3DGS and 2DGS. We utilize two real-world rainy scenarios captured without and with camera flash from the same camera position for compatibility. To analyze the impact of rain on the scene representation we consider occlusion masks with different mask coverage. In general, 3DGS showcases best novel view synthesis quality outperforming NeRF by a large margin making it a robust solution for generating novel views in rainy scenarios. It is noteworthy that the rain is not reconstructed in the scene and subsequently not rendered in the novel views due to its dynamic nature as the raindrops and streaks appear at different positions in each image, however influencing the geometry of the features it occludes. Illumination of the rain with camera flash introduces strong specular highlights which are challenging for methods that learn primarily diffuse appearance or that do not model light scattering or Bidirectional Reflectance Distribution Function (BRDF) at raindrop scale, imposing a bottleneck for all methods. Masking the rain leads to cleaner scene reconstruction and renderings. Therefore, we demonstrate radiance field method’s ability to reliably reconstruct the geometry behind rain allowing applicability to Autonomous Vehicles (AV) and Advanced Driver Assistance Systems (ADAS), traffic monitoring, autonomous navigation, surveillance, object detection and recognition and disaster monitoring.

References

Chen, D., Li, H., Ye, W., Wang, Y., Xie, W., Zhai, S., Wang, N., Liu, H., Bao, H., Zhang, G., 2024a. Pgsr: Planar-based gaussian splatting for efficient and high-fidelity surface reconstruction. *IEEE Transactions on Visualization and Computer Graphics*.

Chen, W.-T., Yifan, W., Kuo, S.-Y., Wetzstein, G., 2024b. Dehazenerf: Multi-image haze removal and 3d shape reconstruction using neural radiance fields. *2024 International Conference on 3D Vision (3DV)*, IEEE, 247–256.

Hore, A., Ziou, D., 2010. Image quality metrics: Psnr vs. ssim. *2010 20th international conference on pattern recognition*, IEEE, 2366–2369.

Huang, B., Yu, Z., Chen, A., Geiger, A., Gao, S., 2024. 2d gaussian splatting for geometrically accurate radiance fields. *ACM SIGGRAPH 2024 conference papers*, 1–11.

Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G., 2023. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM Trans. Graph.*, 42(4), 139–1.

Kulhanek, J., Peng, S., Kukelova, Z., Pollefeys, M., Sattler, T., 2024. WildGaussians: 3D Gaussian Splatting in the Wild. *arXiv preprint arXiv:2407.08447*.

Li, R., Gao, H., Tancik, M., Kanazawa, A., 2023. Nerfacc: Efficient sampling accelerates nerfs. *Proceedings of the IEEE/CVF international conference on computer vision*, 18537–18546.

Li, R., Tancik, M., Kanazawa, A., 2022. Nerfacc: A general nerf acceleration toolbox. *arXiv preprint arXiv:2210.04847*.

Li, Y., Wu, J., Zhao, L., Liu, P., 2024. Derainnerf: 3d scene estimation with adhesive waterdrop removal. *2024 IEEE Intern. Conf. on Robotics and Automation (ICRA)*, IEEE, 2787–2793.

Liu, S., Chen, X., Chen, H., Xu, Q., Li, M., 2024. DeRainGS: Gaussian Splatting for Enhanced Scene Reconstruction in Rainy Environments. *arXiv preprint arXiv:2408.11540*.

Liu, S., Tang, X., Li, Z., He, Y., Ye, C., Liu, J., Huang, B., Zhou, S., Wu, X., 2025. OccluGaussian: Occlusion-Aware Gaussian Splatting for Large Scene Reconstruction and Rendering. *arXiv preprint arXiv:2503.16177*.

Lyu, X., Liu, H., Hou, J., 2024. Rainyscape: Unsupervised rainy scene reconstruction using decoupled neural rendering. *Proceedings of the 32nd ACM International Conference on Multimedia*, 10920–10929.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., Ng, R., 2021. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1), 99–106.

Petrovska, I., Jutzi, B., 2024. Vision through Obstacles-3D Geometric Reconstruction and Evaluation of Neural Radiance Fields (NeRFs). *Remote Sensing*, 16(7), 1188.

Petrovska, I., Jutzi, B., 2025. Impact of Rain on 3D Reconstruction with Multi-View Stereo, Neural Radiance Fields and Gaussian Splatting. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 10, 169–176.

Qian, C., Guo, Y., Li, W., Markkula, G., 2024. WeatherGS: 3D Scene Reconstruction in Adverse Weather Conditions via Gaussian Splatting. *arXiv preprint arXiv:2412.18862*.

Ramazzina, A., Bijelic, M., Walz, S., Sanvito, A., Scheuble, D., Heide, F., 2023. Scatternerf: Seeing through fog with physically-based inverse neural rendering. *Proceedings IEEE/CVF Intern. Conf. on Computer Vision*, 17957–17968.

Sabour, S., Goli, L., Kopanas, G., Matthews, M., Lagun, D., Guibas, L., Jacobson, A., Fleet, D. J., Tagliasacchi, A., 2024. Spotlesssplats: Ignoring distractors in 3d gaussian splatting. *arXiv preprint arXiv:2406.20055*.

Sabour, S., Vora, S., Duckworth, D., Krasin, I., Fleet, D. J., Tagliasacchi, A., 2023. Robustnerf: Ignoring distractors with robust losses. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 20626–20636.

Schonberger, J. L., Frahm, J.-M., 2016. Structure-from-motion revisited. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4104–4113.

Schönberger, J. L., Zheng, E., Frahm, J.-M., Pollefeys, M., 2016. Pixelwise view selection for unstructured multi-view stereo. *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III 14*, Springer, 501–518.

Shi, J., Luthra, A., Li, J., Gao, X., Song, X., Lin, Z., Gu, D., Yu, H., 2025. Occludenerf: Geometry-aware 3d scene inpainting with collaborative score distillation in nerf. *Proceedings of the Computer Vision and Pattern Recognition Conference*, 284–294.

Tancik, M., Weber, E., Ng, E., Li, R., Yi, B., Wang, T., Kristoffersen, A., Austin, J., Salahi, K., Ahuja, A. et al., 2023. Nerfstudio: A modular framework for neural radiance field development. *ACM SIGGRAPH 2023 conference proceedings*, 1–12.

Teng, Z., Lin, B., Chen, T., Yuan, Z., Li, X., Zhang, X., Zhang, S., 2025. Raindrop GS: A benchmark for 3d gaussian splatting under raindrop conditions. *Submitted to The Fourteenth International Conference on Learning Representations*.

Wang, Z., Bovik, A. C., 2009. Mean squared error: Love it or leave it? *IEEE Signal Processing Magazine*, 26(1), 98–117.

Wang, Z., Bovik, A. C., Sheikh, H. R., Simoncelli, E. P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE trans. on image processing*, 13(4), 600–612.

Xu, C., Kerr, J., Kanazawa, A., 2024. Splatfacto-W: A Nerfstudio Implementation of Gaussian Splatting for Unconstrained Photo Collections. *arXiv preprint arXiv:2407.12306*.

Yang, Y., Zhou, Y., Huang, H., 2025. Introducing Unbiased Depth into 2D Gaussian Splatting for High-accuracy Surface Reconstruction. *arXiv preprint arXiv:2503.06587*.

Ye, V., Li, R., Kerr, J., Turkulainen, M., Yi, B., Pan, Z., Seiskari, O., Ye, J., Hu, J., Tancik, M. et al., 2025. gsplat: An open-source library for Gaussian splatting. *Journal of Machine Learning Research*, 26(34), 1–17.

Zhang, R., Isola, P., Efros, A. A., Shechtman, E., Wang, O., 2018. The unreasonable effectiveness of deep features as a perceptual metric. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 586–595.

Zhu, C., Wan, R., Tang, Y., Shi, B., 2023. Occlusion-free scene recovery via neural radiance fields. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 20722–20731.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-1-2026-17-2026