RESPLAT: DEGRADATION-AGNOSTIC FEED-FORWARD GAUSSIAN SPLATTING VIA SELF-GUIDED RESIDUAL DIFFUSION

Anonymous authors
Paper under double-blind review

ABSTRACT

Recent advances in novel view synthesis (NVS) have predominantly focused on ideal, clear input settings, limiting their applicability in real-world environments with common degradations such as blur, low-light, haze, rain, and snow. While some approaches address NVS under specific degradation types, they are often tailored to narrow cases, lacking the generalizability needed for broader scenarios. To address this issue, we propose Restoration-based feed-forward Gaussian Splatting, named *ReSplat*, a novel framework capable of handling degraded multi-view inputs. Our model jointly estimates restored images and gaussians to represent the clear scene for NVS. We enable multi-view consistent universal image restoration by utilizing the 3d gaussians generated during the diffusion sampling process as self-guidance. This results in sharper and more reliable novel views. Notably, our framework adapts to various degradations without prior knowledge of their specific types. Extensive experiments demonstrate that ReSplat significantly outperforms existing methods across challenging conditions, including blur, low-light, haze, rain, and snow, delivering superior quality and robust NVS performance.

1 INTRODUCTION

Novel View Synthesis (NVS) is a task aimed at generating novel views of a scene from a known set of views. NVS strives to accurately estimate the geometry and appearance of a scene, enabling the rendering of realistic images from unseen perspectives. In recent years, Neural Radiance Fields (NeRF) Mildenhall et al. (2021) have revolutionized NVS by utilizing neural networks to represent scenes in a continuous volumetric format, producing highly realistic results. However, NeRF's slow rendering speed has limited its practicality, especially in real-time applications. Solutions like InstantNGP Müller et al. (2022) and TensoRF Chen et al. (2022) have addressed these speed limitations, and Gaussian Splatting Kerbl et al. (2023), introduced later, further accelerated the rendering process. By representing scenes with Gaussian ellipsoids instead of dense point samples like NeRF, Gaussian Splatting maintains competitive visual quality while enabling faster rendering.

Despite the impressive results of NeRF and Gaussian Splatting, generalizable approaches have become a major focus area. Generalizable NeRF aims to synthesize new views without retraining on each new scene, enhancing model flexibility across diverse datasets Wang et al. (2021); Yu et al. (2021b); Wang et al. (2022b); Suhail et al. (2022). Similarly, generalizable Gaussian Splatting extends this concept, offering a faster and adaptable solution for unseen scenes Charatan et al. (2023); Chen et al. (2025); Liu et al. (2025); Ziwen et al. (2024). However, these methods have primarily been developed to work on clean multi-view images captured from controlled environments.

Against this backdrop, scene reconstruction using corrupted images has gained attention. Some studies Ma et al. (2022); Wang et al. (2022a); Yoon & Yoon (2023); Wang et al. (2023); Chen et al. (2023b) are designed to handle specific types of corruption. GAURA Gupta et al. (2024), on the other hand, leverages the capacity of feed-forward NVS models to be pre-trained on large multi-view datasets, proposing a generalizable NeRF model that operates under a variety of degradations. However, GAURA excludes the image restoration capabilities developed in the 2D domain, which limits its performance potential.

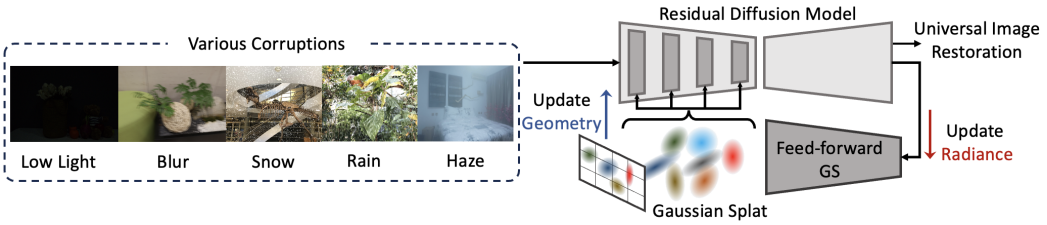

Figure 1: Proposed degradation-agnostic feed-forward Gaussian Splatting (GS) framework. Our framework achieves high-performance universal image restoration and novel view synthesis results through mutual information exchange between the universal image restoration model and the generalizable GS model.

Simply adopting an image restoration model does not fully address this limitation. Universal image restoration is a severely ill-posed problem, with countless possible solutions. This has led to various methods using denoising diffusion models—prominent examples of stochastic models—to address image restoration Fei et al. (2023); Özdenizci & Legenstein (2023). Research has shown that training to predict residual images enables effective image restoration Zhang et al. (2017; 2018); Zamir et al. (2021); Anwar & Barnes (2020), enhancing performance through diffusion-based residual learning.

In this paper, we propose a new generalizable gaussian splatting framework, **ReSplat**, aimed at degradation-agnostic novel view synthesis. At the core of our framework is a method that leverages the model priors of a diffusion-based unified image restoration network through Gaussian splatting. Unlike NeRF's representation, Gaussian splatting uses a point-based representation that enables explicit scene geometry extraction during training. As shown in Fig. 1, generalizable Gaussian splatting models Charatan et al. (2023); Chen et al. (2025); Liu et al. (2025) inherently estimate Gaussian centroids (geometry) using multi-view stereo (MVS) and radiance (color) through multi-view image aggregation. In our framework, a diffusion model iteratively estimates Gaussian centroids, or 3D geometry, leveraging this information to achieve 3D-consistent image restoration.

Our framework specifically adapts a 3D cross-attention module to the residual diffusion model, enabling it to utilize the location information of Gaussian centroids. Here, Gaussian centroids are derived from the point clouds of restored images estimated in the previous time-step. Second, our model performs multi-view aligned pre-filtering when generating Gaussian ellipsoids. This process involves calculating a weight map that is applied to the image features used to generate the Gaussian ellipsoids, helping to achieve artifact-free novel view synthesis. Through these techniques, our model retains the advantages of a generalizable method that operates without a scene optimization process, working effectively even in sparse-view settings while remaining degradation-agnostic. This makes it a more practical NVS model, demonstrating superior NVS and image restoration performance in multiple degradation settings compared to other approaches.

In summary, our contributions are summarized as follows:

1. We propose *ReSplat*, a novel framework for multi-view image restoration using 3DGS.
2. We introduce a multi-view aligned denoising diffusion model for universal image restoration.
3. Our method outperforms other methods in novel view synthesis and image restoration tasks.

2 Related Works

2.1 Generalizable Radiance Fields

Generating realistic images has been a central research topic for many years. Neural scene representations, such as Neural Radiance Fields (NeRF) Mildenhall et al. (2021), have emerged as effective solutions for view synthesis, achieving remarkable results. Subsequent NeRF-based approaches have further improved rendering quality Roessle et al. (2022); Wei et al. (2021); Deng et al. (2022), as well as optimization and rendering speed Sun et al. (2022); Chen et al. (2022); Fridovich-Keil et al. (2022); Yu et al. (2021a); Müller et al. (2022). However, NeRF still requires optimization for each new scene to synthesize novel views. To address this, various studies have proposed generalizable NeRF models Yu et al. (2021b); Wang et al. (2021); Liu et al. (2022); Wang et al. (2022b); Suhail et al. (2022); Cao et al. (2022), enabling cross-scene generalization by learning a view in-

terpolation function from source images. In these generalizable NeRFs, a common technique is to apply volume rendering for aggregating information from images, such as deep features, depth maps, or cost volumes Wang et al. (2021); Liu et al. (2022); Johari et al. (2022); Chen et al. (2021); Xu et al. (2023). GPNR Suhail et al. (2022) and GNT Wang et al. (2022b) utilize transformers to aggregate features, enhancing information interaction along a ray to directly predict RGB values for each pixel. PixelSplat Charatan et al. (2023) and MVSplat Chen et al. (2025) propose generalizable volume rendering techniques that utilize scene parameterization with 3D Gaussian primitives Kerbl et al. (2023). We note that existing generalizable radiance fields have predominantly been studied on clean images, and we aim to address this limitation by developing a universal model utilizing a residual diffusion model.

2.2 Novel View Synthesis with degradations

Some research has advanced novel view synthesis (NVS) by leveraging radiance fields with physics-based multi-view geometry techniques, targeting cases where train-view images require enhancement. NeRF-W Martin-Brualla et al. (2021) addresses variations in illumination and transient occlusions by relaxing strict assumptions on consistency across inputs. Deblur-NeRF Ma et al. (2022) introduces a spatially-varying blur kernel model to handle blurry inputs effectively. RawNeRF Mildenhall et al. (2022) facilitates high-dynamic range (HDR) view synthesis by training NeRF on raw input data and generating raw-format outputs. Similarly, HDR-NeRF Huang et al. (2022) supports exposure control and HDR image synthesis by learning two distinct implicit functions: one for the radiance field and another for tone mapping. LLNeRF Wang et al. (2023) and Aleth-NeRF Cui et al. (2024) conducted research on novel view synthesis under low-light conditions. More recently, DiET-GS Lee & Lee (2025) and DiSR-NeRF Lee et al. (2024) leverage diffusion priors to improve 3D representations from degraded inputs, but they are designed for specific corruption types such as motion blur or low resolution. HQGS Lin et al. and RobustGS Wu et al. (2025) further study Gaussian Splatting under various degraded conditions and propose task-specific 3DGS pipelines to boost robustness in these scenarios. Overall, these studies do not explicitly leverage a pretrained universal image restoration model and remain tailored to specific degradation regimes or 3D configurations. In contrast, we target a degradation-agnostic framework that actively uses a pretrained universal restoration prior within a feed-forward 3DGS pipeline, so that a single model can handle diverse and mixed degradations.

2.3 Universal Image Restoration

Developing a unified model capable of handling multiple degradations has become a growing area of interest. AiRnet Li et al. (2022) introduces a module to align various distributions into a shared distribution using contrastive learning, though this approach can be challenging to train and may limit performance. IDR Zhang et al. (2023) identifies that distinct degradation types can be separated using singular value decomposition (SVD), allowing for clean image reconstruction through reformulation of singular values and vectors. PromptIR Potlapalli et al. (2024) enhances performance by employing a prompt block to capture degradation-specific features. multi-task DINO-based restoration Lin et al. (2023) and mask-based blind restoration Qin et al. (2024) exploit strong visual priors from foundation models. Adair Cui et al. (2025) and Perceive-IR Zhang et al. (2025) further improve all-in-one restoration by adaptively modeling degradation-specific frequency cues and enhancing degradation perception, respectively. Methods such as ProRes Ma et al. (2023) and DA-CLIP Luo et al. (2023) leverage prompt learning to fully utilize the power of large-scale models. Daclip-IR Luo et al. (2024) incorporates a CLIP-based encoder to identify degradation types, extracting semantic information from distorted images to guide a diffusion model in generating high-quality outputs. DiffUIR Zheng et al. (2024) introduces selective hourglass mapping to adapt residual denoising diffusion models Liu et al. (2024) as a comprehensive image restoration approach.

3 Methods

In this section, we provide an overview of the residual denoising diffusion model (RDDM), universal RDDM, and 3D Gaussian splatting as preliminaries. Next, we propose an overall framework for degradation-agnostic feed-forward Gaussian splatting. Additionally, we detail two modules specifically designed to enhance NVS performance.

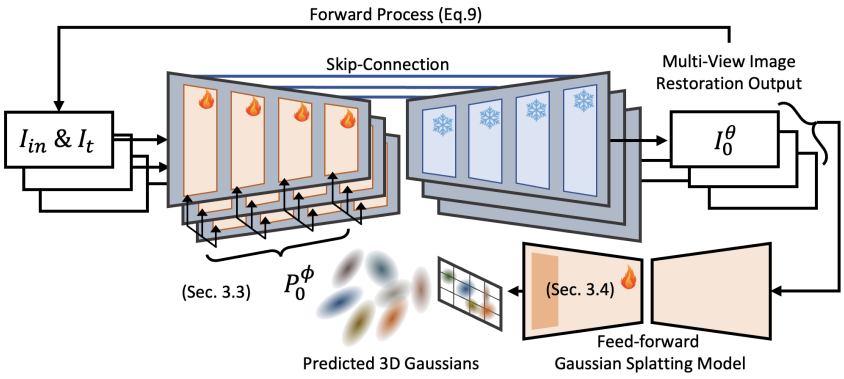

Figure 2: The overall framework for degaradation-agnostic feed-forward gaussian splatting (GS). A diffusion-based image restoration model restores the original image by iteratively estimating the residual image. During this process, feed-forward GS is performed using the original image generated in the intermediate stages of diffusion sampling. By utilizing the Gaussian points information obtained in this process, the diffusion model receives multi-view information in the next diffusion step, enabling more accurate image restoration.

3.1 PRELIMINARIES

3D Gaussian Splatting 3D-GS Kerbl et al. (2023) models a scene using a collection of anisotropic 3D Gaussians, which retain the differential characteristics of volumetric representations while enabling efficient rendering through a tile-based rasterization approach. Beginning with points derived from Structure-from-Motion (SfM), each point serves as the position (mean) μ of a 3D gaussian elipsoids.

$$G(x) = e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)} \tag{1}$$

where x represents a specific point in the 3D scene, and Σ is the covariance matrix of the 3D Gaussian. Σ is constructed from a scaling matrix S and a rotation matrix R with the equation $\Sigma = RSS^TR^T$. For performing tile-based rasterization, the 3D Gaussians $G(x)$ are projected onto the image plane as 2D Gaussians $G'(x)$. The rasterizer then sorts these 2D Gaussians and applies alpha blending:

$$C(x') = \sum_{i\in N} c_i\sigma_i \prod_{j=1}^{i-1}(1-\sigma_j), \quad \sigma_i = \alpha_i G_i'(x') \tag{2}$$

x' represents the queried pixel position, and N denotes the number of sorted 2D Gaussians associated with that pixel.

Feed-forward 3D Gaussian Splatting While vanilla 3DGS optimizes Gaussian parameters per scene, recent feed-forward 3DGS models Liu et al. (2025) predict them in a single forward pass from a few posed views. Given N input images $\{I_{\text{in}}^i\}_{i=1}^N$ and their camera poses $\{\Pi^i\}_{i=1}^N$, the network ϕ maps multi-view features to per-pixel Gaussian primitives:

$$\phi : \{(I_{\text{in}}^i, \Pi^i)\}_{i=1}^N \longmapsto \{(\mu_j, \Sigma_j, \alpha_j, c_j)\}_{j=1}^{H\times W\times N}, \tag{3}$$

where $(\mu_j, \Sigma_j, \alpha_j, c_j)$ denote the center, covariance, opacity, and color of candidate Gaussians. These predictions are then pruned and merged into the explicit Gaussian set P_ϕ^0, which is rendered using the standard splatting formulation above. In addition, the feed-forward 3DGS predicts per-view aggregation weights W^i that are used to combine warped multi-view features at each novel-view pixel; in Sec. 3.4, we modulate these weights with our degradation-aware pre-filtering module.

Residual Denoising Diffusion Model RDDM Liu et al. (2024) uses a standard T-step diffusion model that includes both a forward and a reverse process. In the forward process, one-step noising is formulated as a Markov chain:

$$q(I_t|I_{t-1}, I_{res}) = \mathcal{N}(I_t; I_{t-1} + \alpha_t I_{res}, \beta_t^2\mathbf{I}) \tag{4}$$

where α_t and β_t are the noise coefficients for I_{res} and gaussian noise. I_t is the result at timestep t, and I_{res} represents the residual between the degraded image I_{in} and the clean image I_0, with $I_{\text{res}} =$

Algorithm 1 Training

Input: Clean image, Degraded image: I_0, I_{in};
GT novel view image : I_{nv};
GT residual map: $I_{\text{res}} = I_{\text{in}} - I_0$;
Image Restoration Model: $\theta(*)$;
Feed-forward GS Model: $\phi(*)$;

repeat
  $I_0 \sim q(I_0)$;
  $P_0^{\phi} = \phi(I_0, I_{in})$;
  $t \sim \text{Uniform}(1, \ldots, T)$;
  $\epsilon \sim \mathcal{N}(0, I)$;
  $I_t = I_0 + \bar{\alpha}_t I_{\text{res}} + \bar{\beta}_t \epsilon - \bar{\delta}_t I_{\text{in}}$;
  Take the gradient descent step on

$$\nabla_\theta \| I_{\text{res}} - I^{\theta}_{\text{res}}(P_0^{\phi}, I_t, I_{\text{in}}, t) \|_1 +$$
$$\nabla_\phi \| I_{nv} - I^{\phi}_{\text{ren}}(I_{\text{in}} - I^{\theta}_{\text{res}}, I_{\text{in}}) \|_1;$$

until converged

Algorithm 2 Sampling

Require: Degraded image: I_{in};
Image Restoration Model: $\theta(*)$;
Feed-Forward GS Model: $\phi(*)$;

$\epsilon \sim \mathcal{N}(0, I)$;
$I_T = (1 - \bar{\delta}_T) I_{in} + \bar{\beta}_T \epsilon$;
$P_0^{\phi} = None$;
for $t = T, \ldots, 1$ **do**
  if $t > 1$ **then**
    $I_{t-1} = I_t - \alpha_t I^{\theta}_{res}(P_0^{\phi}, I_t, I_{in}, t) + \delta_t I_{in}$;
    $P_0^{\phi} = \phi(I_{\text{in}} - I^{\theta}_{\text{res}}, I_{in})$;
  else
    $I_{t-1} = I_{in} - I^{\theta}_{res}(P_0^{\phi}, I_t, I_{in}, t)$;
    $I_{nv} = I^{\phi}_{\text{ren}}(I_{\text{in}} - I^{\theta}_{\text{res}}, I_{\text{in}})$;
return I_0, I_{nv}

$I_{\text{in}} - I_0$. In the reverse process, RDDM approximates the true generative distribution $p_\theta(I_{t-1}|I_t)$ by using $q(I_{t-1}|I_t, I_0, I_{\text{res}})$, which is also formulated as a Markov chain when deterministic implicit sampling using DDIM Song et al. (2020):

$$p_\theta(I_{t-1}|I_t) = \mathcal{N}\left(I_{t-1}; I_0^{\theta} + \overline{\alpha}_{t-1} I^{\theta}_{res} + \overline{\beta}_{t-1} \epsilon^{\theta}, 0 \cdot \mathbf{I}\right) \tag{5}$$

In summary, the relatios between I_t and I_{t-1} in both the forward and reverse processes is as follows:

$$I_t = I_{t-1} + \alpha_t I_{res} + \beta_t \epsilon_{t-1} \tag{6}$$

$$I_{t-1} = I_t - (\bar{\alpha}_t - \bar{\alpha}_{t-1}) I^{\theta}_{\text{res}} - (\bar{\beta}_t - \bar{\beta}_{t-1}) \epsilon^{\theta} \tag{7}$$

Universal Residual Denoising Diffusion Model DiffUIR Zheng et al. (2024) utilizes the conditioning mechanism from RDDM and incorporate a shared distribution term (SDT) within the diffusion algorithm for universal image restoration. They adjust the forward process as follows:

$$I_t = I_{t-1} + \alpha_t I_{res} + \beta_t \epsilon_{t-1} - \delta_t I_{in} \tag{8}$$

where $\delta_t I_{\text{in}}$ represents the SDT, and δ is the shared distribution coefficient. The reverse process is also as follows:

$$I_{t-1} = I_t - \alpha_t I^{\theta}_{\text{res}} + \delta_t I_{\text{in}} - (\beta_t^2 / \bar{\beta}_t) \epsilon^{\theta} \tag{9}$$

Finally, in the deterministic implicit sampling process, I_{t-1} and the pseudo clean image I_0^{θ} can be derived using the following equation:

$$I_{t-1} = I_0^{\theta} + \overline{\alpha}_{t-1} I^{\theta}_{res} - \overline{\delta}_{t-1} I_{in} \;\; \text{s.t.} \;\; I_0^{\theta} = I_{in} - I^{\theta}_{res} \tag{10}$$

3.2 Overall Framework

We aim to develop a novel view synthesis model that can be performed under arbitrary degradation. Most of the existing novel view synthesis studies have been conducted on clean images without corruption, and even in the case of studies on situations with corruption, models specialized for specific degradation types are being developed Ma et al. (2022); Wang et al. (2022a); Yoon & Yoon (2023); Wang et al. (2023); Chen et al. (2023b). These studies solve the problem by simultaneously optimizing scene optimization and physical characteristics by implementing physical characteristics that cause specific degradation as a rendering process. Therefore, there is a need for new degradation-agnostic novel view synthesis (NVS) studies. To address this, we propose a new framework, ***ReSplat***, the NVS model that leverages the diffusion prior studied in the field of 2D image restoration.

Training stage As shown in Fig. 2, we combine the feed-forward gaussian splatting (GS) model and the universal image restoration (UIR) model. The GS and UIR models support complementary roles. Unlike NeRF, feed-forward GS inevitably performs multi-view stereo (MVS) within the model because it needs to explicitly extract point clouds. This enables acquisition of 3D scene geometry

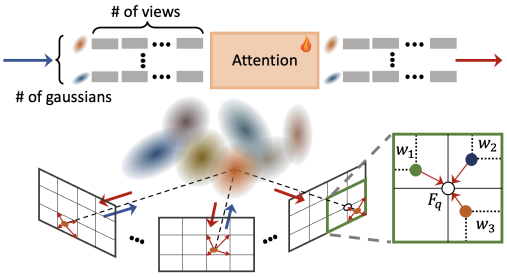

Figure 3: **GS-guided multi-view alignment**. Module embedded in the residual diffusion model that shares info between adjacent views using Gaussian centers.

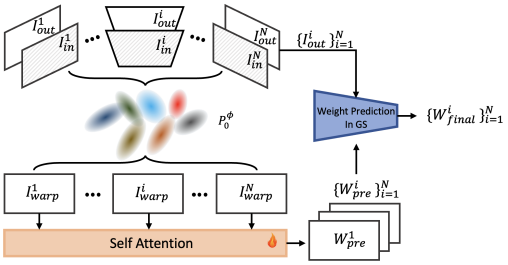

Figure 4: **Pre-filtering with warped features**. Warped inputs are self-attended to form pre-filtering weights for feature aggregation.

information and helps the UIR model find corresponding points for adjacent multi-view images. Meanwhile, the UIR model performs degradation-agnostic image restoration to help feed-forward GS perform NVS using images with corruption removed. In addition, we adopt DiffUIR Zheng et al. (2024), a diffusion model-based UIR model, to perform iterative scene geometry extraction and iterative image refinement so that the UIR results can be gradually refined. The training process can be found in Algorithm 1. The first term $\left\|I_{\text{res}} - I^{\theta}_{\text{res}}(P^{\phi}_0, I_t, I_{\text{in}}, t)\right\|_1$ corresponds to the universal restoration loss $\mathcal{L}_{\text{UIR}}$, which supervises the residual prediction of the UIR model θ. The second term $\left\|I_{nv} - I^{\phi}_{\text{ren}}(I_{\text{in}} - I^{\theta}_{\text{res}}, I_{\text{in}})\right\|_1$ defines the novel-view rendering loss $\mathcal{L}_{\text{NV}}$, which supervises the feed-forward GS model ϕ using the ground-truth clean novel view I_{nv}.

Sampling stage The specific sampling process of ReSplat is shown in Algorithm 2. Given N multi-view input images $\{I^i_{in}\}^N_{i=1}$, the t^{th} noise images $\{I^i_t\}^N_{i=1}$ are generated according to the forward process of the DiffUIR. We generate predicted clean images $\{I^i_0\}^N_{i=1}$ from predicted $\{I^i_{res}\}^N_{i=1}$ by the UIR model. The predicted clean images are used to generate explicit point cloud P^{ϕ}_0 by the MVS module of the feed-forward GS model. Meanwhile, we generate $\{I^i_{t-1}\}^N_{i=1}$ to perform the next diffusion step. After that, we perform a 3d aligned diffusion reverse process using P^{ϕ}_0 (sec. 3.3). Through this, we regenerate the refined $\{I^i_0\}^N_{i=1}$ and P^{ϕ}_0. We repeat the process and perform the feed-forward GS overall process using the finally generated $\{I^i_0\}^N_{i=1}$. In this process, we perform a feature pre-filtering process conditioned on the original corrupted images $\{I^i_{in}\}^N_{i=1}$ to remove points where artifacts exist before the multi-view feature aggregation process, thereby generating a more robust GS output (sec. 3.4). Through this, we can obtain a rendered output for the novel view point.

3.3 GS GUIDED MULTI-VIEW ALIGNMENT

Since the original UIR model is designed for a single image, it is necessary to design a module for enabling multi-view image interaction. As shown in Fig. 3, we adapt a module that performs feature attention in space to UIR by utilizing P^{ϕ}_0, a pseudo geometry generated during the sampling process. Multi-view features are projected toward each gaussian center in P^{ϕ}_0. When there are N multi-view feature vectors $\{f^j_i\}^N_{j=1}$ for the i^{th} center point p_i, we perform self-attention between the corresponding vectors. This process is repeated in the encoder of the diffusion model and helps ensure the 3D consistency of multi-view images. The processed feature vector $f^j_{i,rep}$ is reprojected to the original pixel coordinates. However, since the reprojected point is located in continuous coordinates, not discrete coordinates, it is necessary to propagate to the surrounding discrete coordinates. Therefore, we perform a weighted sum by applying 2D interpolation weights $\{w_i\}$ to all reprojected points existing between adjacent pixels. Each weight is determined by the area of the opposite region, ensuring that features closer to the query point have a higher influence. Therefore, when there is a discrete point q, the multi-view feature F_q that q obtains is as follows.

$$F_q = \sum_i w_i f^j_{i,rep} \quad where \quad i \in Q \tag{11}$$

and Q is the set of the index of all points that exist within the smallest rectangle surrounding the point q.

Table 1: Novel View Synthesis (NV) and Image Restoration (IR) results of five corruption types on LLFF degradation dataset with **three multi-view** inputs. The best scores and second best scores are highlighted with their respective colors.

Method	Operation	Year	Corruption Type	Novel View Synthesis			Multi-View Image Restoration		
				PSNR(↑)	SSIM(↑)	LPIPS(↓)	PSNR(↑)	SSIM(↑)	LPIPS(↓)
AiRnet	IR → NV	2022	Motion Blur	20.11	0.6896	0.4250	21.99	0.7543	0.3769
PromptIR	IR → NV	2023		20.04	0.6872	0.4208	22.14	0.7526	0.3668
GAURA	Only NV	2024		21.28	0.7198	0.4343	-	-	-
DiffUIR	IR → NV	2024		22.75	0.7824	0.3209	26.34	0.8640	0.2951
ReSplat	IR w/ NV	2025		23.15	0.8049	0.3151	27.14	0.8850	0.2713
AiRnet	IR → NV	2022	Snow	20.22	0.6852	0.3026	21.57	0.8184	0.2159
PromptIR	IR → NV	2023		20.54	0.7067	0.2737	23.21	0.8578	0.1912
GAURA	Only NV	2024		20.48	0.7044	0.3195	-	-	-
DiffUIR	IR → NV	2024		24.24	0.8549	0.1826	31.20	0.9627	0.1019
ReSplat	IR w/ NV	2025		24.46	0.8614	0.1677	32.07	0.9685	0.0886
AiRnet	IR → NV	2022	Haze	9.159	0.3841	0.3892	8.871	0.4155	0.2949
PromptIR	IR → NV	2023		9.784	0.4651	0.3508	9.585	0.5251	0.2280
GAURA	Only NV	2024		17.22	0.7205	0.4516	-	-	-
DiffUIR	IR → NV	2024		21.56	0.8392	0.1857	25.57	0.9612	0.0701
ReSplat	IR w/ NV	2025		21.99	0.8471	0.1750	26.45	0.9680	0.0619
AiRnet	IR → NV	2022	Low-light	9.526	0.1364	0.6041	6.388	0.0859	0.7654
PromptIR	IR → NV	2023		6.367	0.0805	0.6240	6.298	0.0805	0.5789
GAURA	Only NV	2024		15.28	0.6627	0.5177	-	-	-
DiffUIR	IR → NV	2024		18.87	0.8241	0.2429	21.88	0.9374	0.1647
ReSplat	IR w/ NV	2025		19.76	0.8276	0.2433	22.82	0.9452	0.1605
AiRnet	IR → NV	2022	Rain	20.49	0.6988	0.3416	23.09	0.8008	0.3018
PromptIR	IR → NV	2023		20.71	0.7175	0.2992	24.78	0.8555	0.2499
GAURA	Only NV	2024		21.78	0.7578	0.4110	-	-	-
DiffUIR	IR → NV	2024		23.51	0.8313	0.2538	29.69	0.9357	0.1919
ReSplat	IR w/ NV	2025		24.11	0.8505	0.2140	31.28	0.9538	0.1533

Table 2: Novel View Synthesis results and multi-view image restoration results of three types (rain+motion blur, snow+motion blur, and haze+snow) on LLFF **mixed degradation** dataset with three multi-view inputs. The best scores are highlighted.

Method	Corruption Type	Novel View Synthesis			Multi-View Image Restoration		
		PSNR(↑)	SSIM(↑)	LPIPS(↓)	PSNR(↑)	SSIM(↑)	LPIPS(↓)
DiffUIR	Rain+Motion Blur	20.07	0.6910	0.4885	20.41	0.7083	0.4653
ReSplat		20.44	0.7090	0.4555	20.74	0.7220	0.4507
DiffUIR	Snow+Motion Blur	21.63	0.7407	0.4076	22.51	0.7757	0.3848
ReSplat		22.00	0.7594	0.3782	22.90	0.7908	0.3661
DiffUIR	Haze+Snow	15.38	0.6978	0.3488	15.52	0.7702	0.2843
ReSplat		20.17	0.7730	0.3148	19.92	0.8067	0.2808

3.4 Pre-filtering with warped features

The final outputs of the UIR, $\{I^i_{\text{out}}\}^N_{i=1}$, are first depth-warped toward the novel pose using P^ϕ_0. The feed-forward GS backbone then produces per-view aggregation weights $\{W^i_{\text{final}}\}^N_{i=1}$ for combining the N warped multi-view features at each novel-view pixel. Since these weights have a critical impact on determining the radiance of the final Gaussian ellipsoids, we introduce a pre-filtering module that is additionally conditioned on the corrupted inputs $\{I^i_{\text{in}}\}^N_{i=1}$.

As illustrated in Fig. 4, the pre-filtering module takes the warped restored and degraded images as input and predicts a per-view reliability map $\{W^i_{\text{pre}}\}^N_{i=1}$, independently of the occlusion-based weights from the GS model. We then modulate the original GS weights by this reliability map:

$$W^i_{\text{final}}(x) = W^i_{\text{pre}}(x) \cdot W^i(x), \tag{12}$$

and use the updated W^i_{final} in the splatting renderer. In other words, the pre-filtering module acts as a soft, degradation-aware gate on top of the standard visibility weights: regions where residual artifacts (e.g., remaining rain streaks, snow blobs, or haze fragments) are strong or inconsistent across views receive lower W^i_{pre} and are down-weighted, while geometry-consistent, clean structures are preserved, leading to a more robust radiance field and improved NVS quality. In practice, we simply replace the original per-view aggregation weights W^i with the updated W^i_{final}. These weights are then used in its standard multi-view feature aggregation and Gaussian rendering pipeline to determine the contribution of each input view at every novel-view location.

4 Experiments

4.1 Experiment Settings

Datasets For training our model, we utilize the synthetic multi-degradation generation pipeline proposed by GAURA Gupta et al. (2024) to construct a multi-view degradation dataset. We use a train-

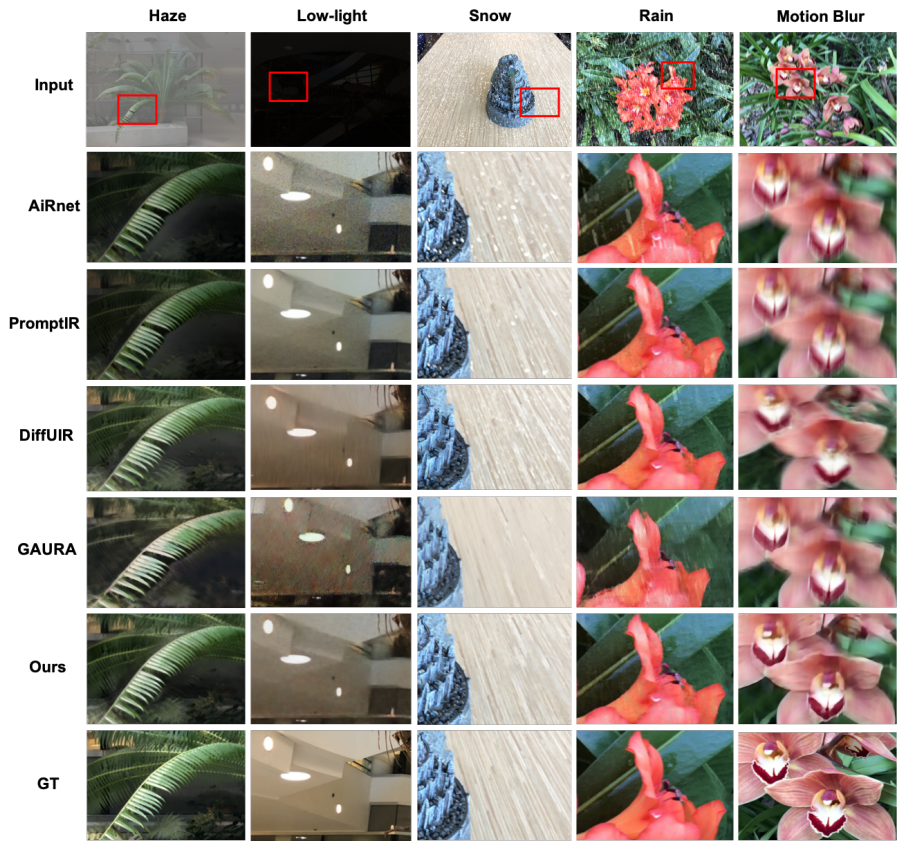

Figure 5: Comparsions of novel view synthesis results of five types (motion blur, snow, haze, low-light, rain) on LLFF degradation dataset.

ing dataset provided by IBRNet Wang et al. (2021), commonly used in novel view synthesis task. The test sets are divided into synthetic and real-world datasets. The synthetic dataset is generated using synthetic degradations applied to the LLFF Mildenhall et al. (2019) dataset. For real-world scenarios, we evaluate our model using the DeblurNeRF Ma et al. (2022) dataset for motion blur, the REVIDE Zhang et al. (2021) dataset for haze, and the LLNeRF Wang et al. (2023) dataset for low-light conditions.

Network We use DiffUIR Zheng et al. (2024), a residual diffusion model, as our baseline for image restoration. We also use MVSGaussian Liu et al. (2025), one of the state-of-the-art models, as the feed-forward GS. To accelerate the training process, MVSGaussian is first trained on our training dataset without image restoration process.

We conduct a comparison of universal image restoration with AiRnet, PromptIR, and DiffUIR. The goal of our approach is to develop an adapter that transforms a UiR model to handle multi-view inputs. Therefore, we utilize a network pretrained with a single-view UiR. For a fair comparison, all models are fine-tuned on our training dataset.

During the inference time, ReSplat uses DDIM sampling strategy with a total of three sampling steps fixed. The inference process for the three multi-view inputs can be completed within one second. For more details, please refer to the supplementary material.

4.2 Quantitative Analysis

Synthetic Degradation As shown in Table 1, we evaluate ReSplat against baselines (AiRnet, PromptIR, DiffUIR, GAURA) across five corruption types: motion blur, snow, haze, low-light, and rain, for both novel view synthesis and multi-view image restoration. Performance is measured using PSNR, SSIM, and LPIPS. For novel view synthesis, ReSplat consistently outperforms other methods, especially in motion blur, snow, and rain scenarios, producing sharper, more perceptually

Table 3: Novel view synthesis results of three types (motion blur, haze, low-light) on real-world degradation datasets.

Method	Type	Novel View Synthesis		
		PSNR(↑)	SSIM(↑)	LPIPS(↓)
AiRnet	Motion Blur	18.59	0.6429	0.4009
PromptIR		18.42	0.6289	0.3959
GAURA		21.54	0.7711	0.3909
DiffUIR		22.76	0.8090	0.2988
ReSplat		22.91	0.8145	0.2922
AiRnet	Haze	15.91	0.7189	0.3290
PromptIR		15.32	0.7106	0.3224
GAURA		16.90	0.8397	0.3920
DiffUIR		17.26	0.8451	0.1900
ReSplat		17.75	0.8511	0.1968
AiRnet	Low-light	9.526	0.1364	0.6040
PromptIR		17.10	0.8322	0.5091
GAURA		19.07	0.8503	0.6301
DiffUIR		22.00	0.8165	0.4958
ReSplat		22.92	0.8578	0.4759

Table 4: Ablation study with Novel View Synthesis results of 5 types (motion blur, snow, haze, low-light, rain) on synthetic degradations. Values represent the average of five degradations. The best scores and second best scores are highlighted.

Model #	Alignment	Pre-Filtering	Novel View Synthesis		
			PSNR(↑)	SSIM(↑)	LPIPS(↓)
1	X	X	22.19	0.8264	0.2372
2	X	O	22.35	0.8290	0.2368
3	O	X	22.46	0.8313	0.2306
4	O	O	22.69	0.8383	0.2230

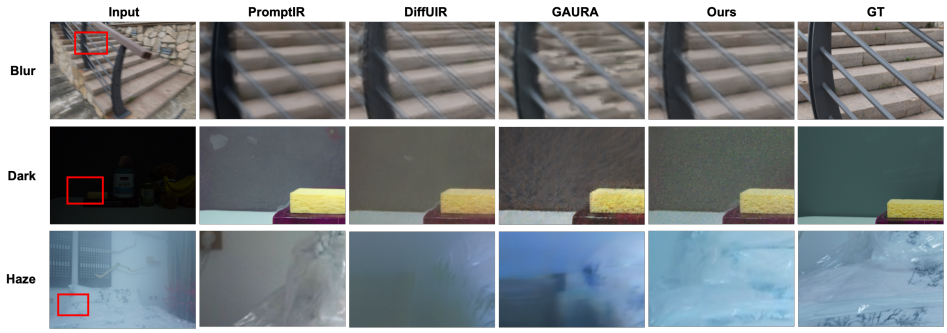

Figure 6: Visual Comparsions of novel view synthesis results of 3 types (motion blur, haze, low-light) on real-world degradation dataset (DeblurNeRF, REVIDE, and LLNeRF dataset).

accurate views. In multi-view image restoration, ReSplat excels in high-corruption cases, achieving the best overall visual fidelity and structural similarity. Notably, it handles complex degradations like heavy rain and motion blur more effectively than competing models, preserving both fine details and global consistency.

Mixed Degradation Table 2 presents a comparison between our method, ReSplat, and the strongest baseline, DiffUIR, under various mixed degradation scenarios, including Rain+Motion Blur, Snow+Motion Blur, and Haze+Snow. ReSplat consistently achieves the best performance across all conditions, significantly outperforming DiffUIR in both novel view synthesis and multi-view image restoration. We conduct a direct comparison between the top-performing method (ReSplat) and the next best (DiffUIR) to highlight the effectiveness of our approach. Notably, in more challenging scenarios such as Snow+Motion Blur and Haze+Snow, ReSplat delivers considerably higher fidelity, as reflected by higher PSNR and SSIM values and lower LPIPS.

Real-World Degradation As shown in Table 3, we evaluate ReSplat against baselines for novel view synthesis under real-world corruptions: motion blur (DeblurNeRF Ma et al. (2022)), haze (REVIDE Zhang et al. (2021)), and low-light (LLNeRF Wang et al. (2023)) dataset. ReSplat achieves the best overall results for motion blur, preserving structural and perceptual quality. In haze, it yields the lowest LPIPS, indicating superior perceptual quality despite similar PSNR/SSIM scores with DiffUIR. Under low-light conditions, ReSplat balances structural integrity and perceptual fidelity, performing competitively across all metrics. In addition, Table 5 and Table 6 report results on in-the-wild rain (NTURain Chen et al. (2018)) and snow (RSVD Chen et al. (2023a)) datasets, where ReSplat consistently outperforms UIR and GS baselines, demonstrating robust generalization.

4.3 Qualitative Comparison and Analysis

Synthetic Degradation Figure 5 shows the results on synthetic degradations including haze, low-light, snow, rain, and motion blur. Our method consistently outperforms competing methods like AiRnet, PromptIR, DiffUIR, and GAURA by producing clearer images with better detail preservation and color balance. Notably, it excels in haze removal, low-light enhancement, and rain streak removal, closely matching the ground truth.

Table 5: Novel View Synthesis (NV) results of **Rain** corruption on the real-world deraining dataset with three multi-view inputs. The best scores and second best scores are highlighted.

Method	Type	PSNR(↑)	SSIM(↑)	LPIPS(↓)
AiRnet	Rain	24.05	0.8183	0.1955
PromptIR		24.23	0.8230	0.1801
GAURA		19.39	0.6602	0.3987
DiffUIR		23.99	0.8145	0.2094
ReSplat		24.35	0.8232	0.1772

Table 6: Novel View Synthesis (NV) results of **Snow** corruption on the real-world desnowing dataset with three multi-view inputs. The best scores and second best scores are highlighted.

Method	Type	PSNR(↑)	SSIM(↑)	LPIPS(↓)
AiRnet	Snow	20.23	0.7035	0.3103
PromptIR		21.27	0.7192	0.3007
GAURA		20.22	0.7578	0.3647
DiffUIR		22.12	0.8215	0.2277
ReSplat		22.45	0.8263	0.2175

Figure 7: Qualitative comparison of the alignment module. The top row shows the restored RGB outputs, while the bottom row visualizes the corresponding error maps.

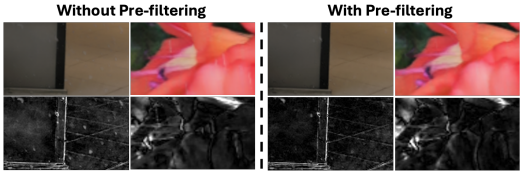

Figure 8: Qualitative comparison of pre-filtering module. The top row shows the restored RGB outputs, while the bottom row visualizes the corresponding error maps.

Real-World Degradation Figure 6 presents results on real-world degradations such as blur, low-light scenes, and haze. Our approach effectively reconstructs details and reduces artifacts, outperforming other methods which often introduce noise or fail to restore fine structures. The results demonstrate that our method achieves superior restoration, preserving natural colors and sharpness in challenging real-world scenarios.

4.4 Ablation Studies

We conduct an ablation study using four model variants: Model 1 as baseline, Model 2 with pre-filtering, Model 3 with alignment, and Model 4 with both. Metrics are averaged over five LLFF degradation datasets. As shown in Table 4, each component contributes to improved performance in novel view synthesis. Adding pre-filtering (Model 2) increases PSNR and reduces LPIPS, indicating a modest improvement in reconstruction quality. Alignment alone (Model 3) further enhances PSNR and other quality metrics compared to the baseline. When both alignment and pre-filtering are applied (Model 4), the model achieves the best overall results, with a PSNR of 22.69, demonstrating a clear effect across various degradations. Qualitative comparisons in Fig. 7 and Fig. 8 illustrate that alignment reduces multi-view geometric inconsistencies, while pre-filtering suppresses residual artifacts without destroying fine structures in the rendered novel views.

5 Limitations

Although ReSplat performs well across various degradation scenarios, several limitations remain. The diffusion-based refinement increases computational cost and memory usage compared to purely feed-forward 3DGS, making efficient high-resolution scaling an important direction for future work. The method also inherits the representational biases of Gaussian Splatting, showing reduced fidelity in scenes with strong transparency, specularities, or pronounced view-dependent effects. Finally, ReSplat depends on a pretrained universal image restoration model for guidance; while this introduces a dependency, it also enables the framework to naturally benefit from future advances in universal restoration without architectural changes.

6 Conclusion

We present a feed-forward gaussian splatting framework for degradation-agnostic novel view synthesis. By integrating a residual diffusion model with 3D cross-attention and multi-view pre-filtering, our method robustly restores images and improves geometry estimation, outperforming existing approaches in both novel view synthesis and universal multi-view image restoration.

REFERENCES

Saeed Anwar and Nick Barnes. Densely residual laplacian super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(3):1192–1204, 2020.

Ang Cao, Chris Rockwell, and Justin Johnson. Fwd: Real-time novel view synthesis with forward warping and depth. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 15713–15724, 2022.

David Charatan, Sizhe Li, Andrea Tagliasacchi, and Vincent Sitzmann. pixelsplat: 3d gaussian splats from image pairs for scalable generalizable 3d reconstruction. *arXiv preprint arXiv:2312.12337*, 2023.

Anpei Chen, Zexiang Xu, Fuqiang Zhao, Xiaoshuai Zhang, Fanbo Xiang, Jingyi Yu, and Hao Su. Mvsnerf: Fast generalizable radiance field reconstruction from multi-view stereo. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 14124–14133, 2021.

Anpei Chen, Zexiang Xu, Andreas Geiger, Jingyi Yu, and Hao Su. Tensorf: Tensorial radiance fields. In *European Conference on Computer Vision*, pp. 333–350. Springer, 2022.

Haoyu Chen, Jingjing Ren, Jinjin Gu, Hongtao Wu, Xuequan Lu, Haoming Cai, and Lei Zhu. Snow removal in video: A new dataset and a novel method. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 13165–13176. IEEE, 2023a.

Jie Chen, Cheen-Hau Tan, Junhui Hou, Lap-Pui Chau, and He Li. Robust video content alignment and compensation for rain removal in a cnn framework. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 6286–6295, 2018.

Wei-Ting Chen, Wang Yifan, Sy-Yen Kuo, and Gordon Wetzstein. Dehazenerf: Multiple image haze removal and 3d shape reconstruction using neural radiance fields. *arXiv preprint arXiv:2303.11364*, 2023b.

Yuedong Chen, Haofei Xu, Chuanxia Zheng, Bohan Zhuang, Marc Pollefeys, Andreas Geiger, Tat-Jen Cham, and Jianfei Cai. Mvsplat: Efficient 3d gaussian splatting from sparse multi-view images. In *European Conference on Computer Vision*, pp. 370–386. Springer, 2025.

Yuning Cui, Syed Waqas Zamir, Salman Khan, Alois Knoll, Mubarak Shah, and Fahad Shahbaz Khan. Adair: Adaptive all-in-one image restoration via frequency mining and modulation. In *13th International Conference on Learning Representations, ICLR 2025*, pp. 57335–57356. International Conference on Learning Representations, ICLR, 2025.

Ziteng Cui, Lin Gu, Xiao Sun, Xianzheng Ma, Yu Qiao, and Tatsuya Harada. Aleth-nerf: Illumination adaptive nerf with concealing field assumption. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 1435–1444, 2024.

Kangle Deng, Andrew Liu, Jun-Yan Zhu, and Deva Ramanan. Depth-supervised nerf: Fewer views and faster training for free. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 12882–12891, 2022.

Ben Fei, Zhaoyang Lyu, Liang Pan, Junzhe Zhang, Weidong Yang, Tianyue Luo, Bo Zhang, and Bo Dai. Generative diffusion prior for unified image restoration and enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9935–9946, 2023.

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*, pp. 5501–5510, 2022.

Vinayak Gupta, Rongali Simhachala Venkata Girish, Ayush Tewari, Kaushik Mitra, et al. Gaura: Generalizable approach for unified restoration and rendering of arbitrary views. *arXiv preprint arXiv:2407.08221*, 2024.

Xin Huang, Qi Zhang, Ying Feng, Hongdong Li, Xuan Wang, and Qing Wang. Hdr-nerf: High dynamic range neural radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 18398–18408, 2022.

Mohammad Mahdi Johari, Yann Lepoittevin, and François Fleuret. Geonerf: Generalizing nerf with geometry priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 18365–18375, 2022.

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4), 2023.

Jie Long Lee, Chen Li, and Gim Hee Lee. Disr-nerf: Diffusion-guided view-consistent super-resolution nerf. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 20561–20570, 2024.

Seungjun Lee and Gim Hee Lee. Diet-gs: Diffusion prior and event stream-assisted motion deblurring 3d gaussian splatting. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 21739–21749, 2025.

Boyun Li, Xiao Liu, Peng Hu, Zhongqin Wu, Jiancheng Lv, and Xi Peng. All-in-one image restoration for unknown corruption. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 17452–17462, 2022.

Xin Lin, Shi Luo, Xiaojun Shan, Xiaoyu Zhou, Chao Ren, Lu Qi, Ming-Hsuan Yang, and Nuno Vasconcelos. Hqgs: High-quality novel view synthesis with gaussian splatting in degraded scenes. In *The Thirteenth International Conference on Learning Representations*.

Xin Lin, Jingtong Yue, Kelvin CK Chan, Lu Qi, Chao Ren, Jinshan Pan, and Ming-Hsuan Yang. Multi-task image restoration guided by robust dino features. *arXiv preprint arXiv:2312.01677*, 2023.

Jiawei Liu, Qiang Wang, Huijie Fan, Yinong Wang, Yandong Tang, and Liangqiong Qu. Residual denoising diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2773–2783, 2024.

Tianqi Liu, Guangcong Wang, Shoukang Hu, Liao Shen, Xinyi Ye, Yuhang Zang, Zhiguo Cao, Wei Li, and Ziwei Liu. Mvsgaussian: Fast generalizable gaussian splatting reconstruction from multi-view stereo. In *European Conference on Computer Vision*, pp. 37–53. Springer, 2025.

Yuan Liu, Sida Peng, Lingjie Liu, Qianqian Wang, Peng Wang, Christian Theobalt, Xiaowei Zhou, and Wenping Wang. Neural rays for occlusion-aware image-based rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7824–7833, 2022.

Ziwei Luo, Fredrik K Gustafsson, Zheng Zhao, Jens Sjölund, and Thomas B Schön. Controlling vision-language models for universal image restoration. *arXiv preprint arXiv:2310.01018*, 3(8), 2023.

Ziwei Luo, Fredrik K Gustafsson, Zheng Zhao, Jens Sjölund, and Thomas B Schön. Photo-realistic image restoration in the wild with controlled vision-language models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6641–6651, 2024.

Jiaqi Ma, Tianheng Cheng, Guoli Wang, Qian Zhang, Xinggang Wang, and Lefei Zhang. Prores: Exploring degradation-aware visual prompt for universal image restoration. *arXiv preprint arXiv:2306.13653*, 2023.

Li Ma, Xiaoyu Li, Jing Liao, Qi Zhang, Xuan Wang, Jue Wang, and Pedro V Sander. Deblur-nerf: Neural radiance fields from blurry images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 12861–12870, 2022.

Ricardo Martin-Brualla, Noha Radwan, Mehdi SM Sajjadi, Jonathan T Barron, Alexey Dosovitskiy, and Daniel Duckworth. Nerf in the wild: Neural radiance fields for unconstrained photo collections. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7210–7219, 2021.

Ben Mildenhall, Pratul P Srinivasan, Rodrigo Ortiz-Cayon, Nima Khademi Kalantari, Ravi Ramamoorthi, Ren Ng, and Abhishek Kar. Local light field fusion: Practical view synthesis with prescriptive sampling guidelines. *ACM Transactions on Graphics (TOG)*, 38(4):1–14, 2019.

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.

Ben Mildenhall, Peter Hedman, Ricardo Martin-Brualla, Pratul P Srinivasan, and Jonathan T Barron. 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*, pp. 16190–16199, 2022.

Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *arXiv preprint arXiv:2201.05989*, 2022.

Ozan Özdenizci and Robert Legenstein. Restoring vision in adverse weather conditions with patch-based denoising diffusion models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(8):10346–10357, 2023.

Vaishnav Potlapalli, Syed Waqas Zamir, Salman H Khan, and Fahad Shahbaz Khan. Promptir: Prompting for all-in-one image restoration. *Advances in Neural Information Processing Systems*, 36, 2024.

Chu-Jie Qin, Rui-Qi Wu, Zikun Liu, Xin Lin, Chun-Le Guo, Hyun Hee Park, and Chongyi Li. Restore anything with masks: Leveraging mask image modeling for blind all-in-one image restoration. In *European Conference on Computer Vision*, pp. 364–380. Springer, 2024.

Barbara Roessle, Jonathan T Barron, Ben Mildenhall, Pratul P Srinivasan, and Matthias Nießner. Dense depth priors for neural radiance fields from sparse input views. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 12892–12901, 2022.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. *arXiv preprint arXiv:2010.02502*, 2020.

Mohammed Suhail, Carlos Esteves, Leonid Sigal, and Ameesh Makadia. Generalizable patch-based neural rendering. In *European Conference on Computer Vision*, pp. 156–174. Springer, 2022.

Cheng Sun, Min Sun, and Hwann-Tzong Chen. Direct voxel grid optimization: Super-fast convergence for radiance fields reconstruction. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5459–5469, 2022.

Chen Wang, Xian Wu, Yuan-Chen Guo, Song-Hai Zhang, Yu-Wing Tai, and Shi-Min Hu. Nerf-sr: High quality neural radiance fields using supersampling. In *Proceedings of the 30th ACM International Conference on Multimedia*, pp. 6445–6454, 2022a.

Haoyuan Wang, Xiaogang Xu, Ke Xu, and Rynson WH Lau. Lighting up nerf via unsupervised decomposition and enhancement. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 12632–12641, 2023.

Peihao Wang, Xuxi Chen, Tianlong Chen, Subhashini Venugopalan, Zhangyang Wang, et al. Is attention all nerf needs? *arXiv preprint arXiv:2207.13298*, 2022b.

Qianqian Wang, Zhicheng Wang, Kyle Genova, Pratul P Srinivasan, Howard Zhou, Jonathan T Barron, Ricardo Martin-Brualla, Noah Snavely, and Thomas Funkhouser. Ibrnet: Learning multi-view image-based rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4690–4699, 2021.

Yi Wei, Shaohui Liu, Yongming Rao, Wang Zhao, Jiwen Lu, and Jie Zhou. Nerfingmvs: Guided optimization of neural radiance fields for indoor multi-view stereo. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 5610–5619, 2021.

Anran Wu, Long Peng, Xin Di, Xueyuan Dai, Chen Wu, Yang Wang, Xueyang Fu, Yang Cao, and Zheng-Jun Zha. Robustgs: Unified boosting of feedforward 3d gaussian splatting under low-quality conditions. *arXiv preprint arXiv:2508.03077*, 2025.

Haofei Xu, Anpei Chen, Yuedong Chen, Christos Sakaridis, Yulun Zhang, Marc Pollefeys, Andreas Geiger, and Fisher Yu. Murf: Multi-baseline radiance fields. *arXiv preprint arXiv:2312.04565*, 2023.

Youngho Yoon and Kuk-Jin Yoon. Cross-guided optimization of radiance fields with multi-view image super-resolution for high-resolution novel view synthesis. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 12428–12438, 2023.

Alex Yu, Ruilong Li, Matthew Tancik, Hao Li, Ren Ng, and Angjoo Kanazawa. Plenoctrees for real-time rendering of neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 5752–5761, 2021a.

Alex Yu, Vickie Ye, Matthew Tancik, and Angjoo Kanazawa. pixelnerf: Neural radiance fields from one or few images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4578–4587, 2021b.

Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 14821–14831, 2021.

Jinghao Zhang, Jie Huang, Mingde Yao, Zizheng Yang, Hu Yu, Man Zhou, and Feng Zhao. Ingredient-oriented multi-degradation learning for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5825–5835, 2023.

Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE transactions on image processing*, 26 (7):3142–3155, 2017.

Xinyi Zhang, Hang Dong, Jinshan Pan, Chao Zhu, Ying Tai, Chengjie Wang, Jilin Li, Feiyue Huang, and Fei Wang. Learning to restore hazy video: A new real-world dataset and a new method. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9239–9248, 2021.

Xu Zhang, Jiaqi Ma, Guoli Wang, Qian Zhang, Huan Zhang, and Lefei Zhang. Perceive-ir: Learning to perceive degradation better for all-in-one image restoration. *IEEE Transactions on Image Processing*, 2025.

Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. Residual dense network for image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2472–2481, 2018.

Dian Zheng, Xiao-Ming Wu, Shuzhou Yang, Jian Zhang, Jian-Fang Hu, and Wei-Shi Zheng. Selective hourglass mapping for universal image restoration based on diffusion model. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 25445–25455, 2024.

Chen Ziwen, Hao Tan, Kai Zhang, Sai Bi, Fujun Luan, Yicong Hong, Li Fuxin, and Zexiang Xu. Long-lrm: Long-sequence large reconstruction model for wide-coverage gaussian splats. *arXiv preprint arXiv:2410.12781*, 2024.