CoReSplat: Consensus-Relative Residual Compensation for Distractor-free 3D Gaussian Splatting

Anonymous authors
Paper under double-blind review

Abstract

Distractor-free 3D Gaussian Splatting (3DGS) aims to recover a static scene from real-world image collections containing dynamic distractors, which introduce inconsistent observations that can corrupt scene optimization and produce rendering artifacts. Existing methods mitigate distractors through reliability-aware supervision or explicit scene decomposition. However, photometric residuals induced by retained observations may still reflect both under-explained scene content and view-specific discrepancies, leaving it unclear how these residuals should guide shared-scene updates. To address this supervision-to-update ambiguity, we propose CoReSplat, a consensus-relative residual compensation framework that leverages cross-view consensus over residual histories to identify disagreement for training-time compensation. Specifically, CoReSplat maintains a per-view residual history to accumulate cross-iteration evidence of residual tendencies. It then aggregates geometrically corresponding residual histories into a cross-view consensus reference that captures their agreement across views. The departure of each view's residual history from this reference defines its consensus-relative disagreement, which is applied as a stop-gradient additive correction to the current prediction before backpropagation. In this way, CoReSplat dynamically reshapes the effective photometric residual relative to the cross-view consensus, while reducing the immediate influence of view-specific discrepancies on scene updates. Experiments on two real-world benchmarks demonstrate improved novel-view synthesis and consistent performance gains when CoReSplat is integrated into representative distractor-free 3DGS baselines.

1 Introduction

3D Gaussian Splatting (3DGS) (Kerbl et al., 2023) provides an efficient explicit scene representation for high-fidelity novel-view synthesis, enabling real-time rendering through differentiable Gaussian rasterization. Conventional 3DGS reconstruction, however, assumes that multi-view observations can be consistently explained by a shared static 3D representation. This assumption is frequently violated in casually captured real-world image collections, where moving people, vehicles, temporary occluders, and other dynamic distractors appear only in subsets of the views (Sabour et al., 2025). When these view-inconsistent observations are absorbed into the reconstructed Gaussian scene, they can produce floaters, ghosting, and other artifacts in novel views (Wang et al., 2025b). Distractor-free 3DGS therefore seeks to recover a clean static scene representation from such inconsistent observations while preserving the geometric and appearance details necessary for faithful novel-view synthesis (Fu et al., 2025b).

Prior work has substantially improved robustness to these observations from several complementary directions. Reliability-aware approaches use robust estimation, uncertainty prediction, pretrained visual features, or learned masks to suppress observations that are unlikely to belong to the persistent scene (Kulhanek et al., 2024; Sabour et al., 2025; Chen et al., 2026). Representation-based methods instead model inconsistent content through dynamic or view-conditioned scene components, separating it from the static scene representation (Dahmani et al., 2024; Zhang et al., 2024; Wang et al., 2025b;a; Lin et al., 2025). More recent methods further intervene in supervision and optimization by regulating training dynamics, recovering multi-view supervision, enforcing cross-model consis-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

tency, or coordinating cross-view gradients (Fu et al., 2025b; Wang et al., 2026; Yang et al., 2026; Kang et al., 2026). Collectively, these methods improve robustness by regulating observation reliability, scene representation, or downstream optimization. Yet once an observation is retained for training, it remains unclear how to translate the photometric discrepancy from that observation into an update of the shared static 3D scene representation.

This question arises because reliability does not fully determine how to attribute updates. A reliability mechanism determines whether, and to what extent, an observation participates in the reconstruction objective. However, the resulting photometric residual need not be attributable to a single source of discrepancy. It may jointly reflect scene content that the current representation still under-explains and view-specific effects from residual distractors, viewpoint-dependent appearance, or other modeling mismatch. Directly fitting the full residual can therefore encourage the shared representation to absorb view-specific variation, whereas uniformly suppressing it may weaken signals needed to recover under-explained scene content. We call this unresolved decision a *supervision-to-update ambiguity*: retaining an observation for training does not uniquely determine how its photometric discrepancy should influence the subsequent update of the shared representation. The central challenge is therefore to preserve residual signals that remain informative for scene recovery while limiting the influence of view-specific discrepancies on the shared representation.

We address this ambiguity directly in residual space and propose CoReSplat, which introduces consensus-relative residual compensation into distractor-free 3DGS. Our key insight is that whether a photometric discrepancy should directly influence the shared representation cannot be determined from its instantaneous value alone; instead, it should be interpreted in the context of evidence accumulated over optimization and across views. To this end, CoReSplat maintains supervision-weighted residual histories as cross-iteration evidence, uses scene geometry to establish cross-view correspondences, and aggregates the aligned histories into a cross-view consensus reference. CoReSplat then defines consensus-relative disagreement as the departure of a view's residual history from this reference. Rather than suppressing the retained residual as a whole, CoReSplat applies only the resulting disagreement as a stop-gradient additive correction to the current prediction. This compensation reshapes the effective photometric residual before backpropagation, guiding scene optimization with a residual signal calibrated relative to cross-view consensus.

Our contributions are summarized as follows:

1. We formulate distractor-free 3DGS from a supervision-to-update perspective, separating observation reliability from the update attribution of photometric discrepancies associated with retained supervision.
2. We construct a cross-view residual consensus from per-view residual histories using geometry-constrained correspondences, establishing a relative reference for interpreting per-view residuals with respect to cross-view agreement.
3. We define consensus-relative disagreement as the deviation of each view's residual history from the cross-view consensus, and use it for residual compensation to reshape the effective photometric residual.

2 Related Work

Neural radiance fields (NeRF) (Mildenhall et al., 2020) established continuous volumetric representations for novel-view synthesis, with later work improving scale handling, unbounded scenes, and optimization efficiency (Barron et al., 2021; 2022; Müller et al., 2022). 3D Gaussian Splatting (3DGS) (Kerbl et al., 2023) introduced an explicit Gaussian representation with efficient differentiable rasterization, followed by advances in anti-aliasing, structured representations, and reconstruction quality (Yu et al., 2024; Lu et al., 2024; Huang et al., 2024). These foundations have since been extended to unconstrained imagery, including casual, sparse, and feed-forward reconstruction settings (Tang et al., 2025; Pan et al., 2026; Furutani et al., 2026).

Reliability-Aware Supervision. These works improve distractor-free reconstruction by estimating observation reliability during optimization. NeRF-W (Martin-Brualla et al., 2021) models transient uncertainty with per-image appearance, RobustNeRF (Sabour et al., 2023) derives robust weights from photometric residuals, and NeRF On-the-go (Ren et al., 2024) incorporates pretrained visual

features. Subsequent 3DGS methods such as SpotLessSplats (Sabour et al., 2025) and WildGaussians (Kulhanek et al., 2024) combine semantic or perceptual cues with learned reliability estimation, while later approaches further exploit geometric or entity priors, object- or patch-level modeling, adaptive uncertainty, and training-dependent residual statistics (Chen et al., 2024; Ungermann et al., 2024; Wang et al., 2024a; Markin et al., 2024; Chen et al., 2026; Hui et al., 2026). These methods use residual, feature, geometric, or semantic evidence to regulate supervision participation. In contrast, CoReSplat operates after supervision has been retained, using residual histories to regulate how the remaining photometric discrepancies affect subsequent scene updates.

Representation-Level Modeling. A complementary line handles view inconsistency by modifying the scene representation. Early radiance-field methods model image-specific appearance or visibility through latent representations and view-dependent components (Chen et al., 2022; Yang et al., 2023; Li et al., 2023). Their Gaussian counterparts, including SWAG (Dahmani et al., 2024), GS-W (Zhang et al., 2024), Splatfacto-W (Xu et al., 2024), and WE-GS (Wang et al., 2024b), encode image-conditioned appearance, opacity, or spherical-harmonic variations in Gaussian attributes. Other methods explicitly separate persistent from non-persistent content: D^{2}NeRF (Wu et al., 2022) decomposes static and time-varying radiance fields, while later 3DGS approaches use shared static and per-view distractor Gaussians (Wang et al., 2025b), static-background and dynamic-foreground components (Wang et al., 2025a), hybrid 3D-static and 2D-transient representations (Lin et al., 2025), or deformable transient fields (Park et al., 2026b). In contrast, we retain a single static scene representation and use its residual histories and cross-view consensus as training-only optimization states rather than trainable components for representing transient content.

Supervision and Optimization Regulation. Recent work extends robust reconstruction beyond reliability estimation and scene decomposition by regulating training supervision and optimization. RobustSplat (Fu et al., 2025b) and RobustSplat++ (Fu et al., 2025a) use delayed densification and stage-aware supervision, while DualSplat (Wang et al., 2026) and progressive filtering approaches (Li & Cheung, 2025; Seo et al., 2026; Liang et al., 2026; Park et al., 2026a) organize reconstruction into successive filtering and refinement stages. Other methods modify supervision itself: RoGSplat (Kong et al., 2025) rectifies unreliable regions with generative priors, whereas PriSplat (Yang et al., 2026) propagates reliable multi-view information into masked regions. At the optimization level, AsymGS (Li et al., 2025) exploits consistency across asymmetric reconstruction paths, multi-view methods coordinate Gaussian updates across views (Du et al., 2026), and HarmoGS (Kang et al., 2026) further addresses conflicting view-specific gradients. Unlike them, CoReSplat operates at the supervision-to-update interface after reliability selection, using consensus-relative disagreement to reshape the photometric residual before gradient formation.

3 METHOD

3.1 PRELIMINARIES

3D Gaussian Splatting. Vanilla 3D Gaussian Splatting (3DGS) (Kerbl et al., 2023) represents a scene as N anisotropic Gaussian primitives $\mathcal{G}_\theta = \{g_i\}_{i=1}^N$, where $g_i = (\boldsymbol{\mu}_i, \mathbf{R}_i, \mathbf{S}_i, \alpha_i, \mathbf{c}_i)$ denotes its center, rotation, scale, opacity, and spherical-harmonic appearance coefficients. Its covariance is parameterized as $\boldsymbol{\Sigma}_i = \mathbf{R}_i \mathbf{S}_i \mathbf{S}_i^\top \mathbf{R}_i^\top$. For a camera π_v, projected Gaussians are depth-sorted and alpha-composited as

$$\widehat{\mathbf{I}}_v(\mathbf{p}) = \sum_{i \in \mathcal{N}(\mathbf{p})} \mathbf{c}_i(\pi_v)\, a_i(\mathbf{p}) \prod_{j<i} (1 - a_j(\mathbf{p}))\,, \tag{1}$$

where $a_i(\mathbf{p})$ denotes the opacity-weighted projected footprint. We write the rendering at iteration t compactly as $\widehat{\mathbf{I}}_v^t = \mathcal{R}(\mathcal{G}_{\theta^t}, \pi_v)$.

Problem Formulation. CoReSplat augments an existing distractor-free 3DGS pipeline while preserving its native scene representation, supervision mechanism, and reconstruction objective. For a training observation $\mathbf{I}_v$, let $\boldsymbol{\omega}_v^t \in [0,1]^{H \times W}$ denote the native pixel-wise supervision weights at iteration t, with $\boldsymbol{\omega}_v^t = \mathbf{1}$ when no explicit weighting is used. The corresponding signed RGB residual is $\mathbf{e}_v^t = \mathbf{I}_v - \widehat{\mathbf{I}}_v^t$. While $\boldsymbol{\omega}_v^t$ determines how strongly each observation participates in reconstruction, it does not determine how the remaining discrepancy should influence the next shared-scene update. As illustrated in Fig. 1, CoReSplat addresses this supervision-to-update ambiguity by contextualizing residuals across iterations and views before they enter scene optimization.

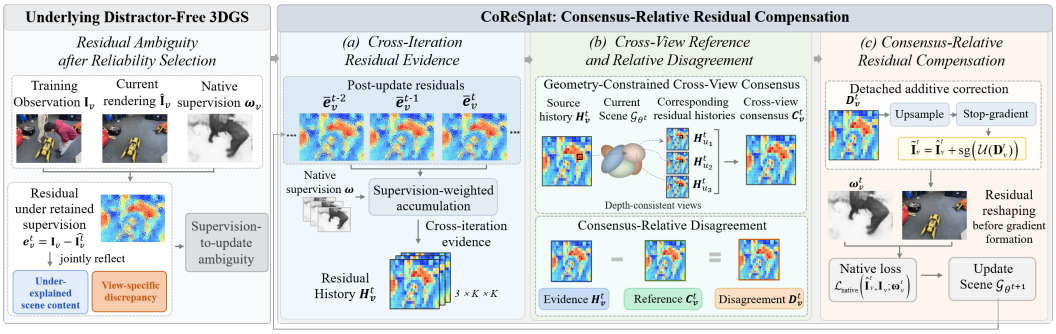

Figure 1: Overview of CoReSplat. After the underlying pipeline determines supervision participation, CoReSplat contextualizes retained residuals across iterations and views. (a) Supervision-weighted post-update residuals form per-view histories. (b) Geometry-constrained correspondence constructs a cross-view consensus, and source history minus consensus yields the consensus-relative disagreement. (c) This disagreement is applied as a stop-gradient correction to the current rendering, reshaping the residual before the native reconstruction objective.

3.2 Cross-Iteration Residual Evidence

An instantaneous photometric residual reflects only a single optimization state and cannot indicate whether a discrepancy persists as the scene evolves. CoReSplat therefore accumulates signed RGB residuals across iterations into per-view histories as cross-iteration evidence.

After updating the Gaussian scene from θ^t to θ^{t+1}, we re-render the sampled view as $\hat{\mathbf{I}}_v^{t+1} = \mathcal{R}(\mathcal{G}_{\theta^{t+1}}, \pi_v)$ and compute the detached post-update residual $\bar{\mathbf{e}}_v^{t+1} = \mathbf{I}_v - \hat{\mathbf{I}}_v^{t+1}$. To align this history with the supervision responsible for the update, we reuse the native pixel-wise weights $\boldsymbol{\omega}_v^t$. The weighted residual and supervision maps are reduced to a $K \times K$ grid by adaptive average pooling $\mathcal{P}_K(\cdot)$, preserving coarse spatial organization for the cross-view correspondence

For each view, we maintain a signed-residual accumulator $\mathbf{A}_v \in \mathbb{R}^{3\times K\times K}$ and a supervision-weight accumulator $\mathbf{B}_v \in \mathbb{R}^{1\times K\times K}$:

$$\mathbf{A}_v^{t+1} = \mathbf{A}_v^t + \mathcal{P}_K\left(\boldsymbol{\omega}_v^t \odot \bar{\mathbf{e}}_v^{t+1}\right), \qquad \mathbf{B}_v^{t+1} = \mathbf{B}_v^t + \mathcal{P}_K\left(\boldsymbol{\omega}_v^t\right), \tag{2}$$

where $\boldsymbol{\omega}_v^t$ is broadcast over RGB channels. The per-view residual history is then defined as

$$\mathbf{H}_v^{t+1} = \mathbf{A}_v^{t+1} \oslash \mathbf{B}_v^{t+1}, \tag{3}$$

for spatial grid locations with positive accumulated supervision weight, where $\oslash$ denotes element-wise division with $\mathbf{B}_v^{t+1}$ broadcast over RGB channels; unsupported spatial grid locations retain their previous history. Thus, $\mathbf{H}_v^{t+1}$ summarizes the supervision-normalized signed residual tendency across iterations, while retaining the discrepancy direction required for subsequent additive compensation.

3.3 Geometry-Constrained Cross-View Residual Consensus

Although $\mathbf{H}_v^t$ provides cross-iteration evidence, it remains defined in view-dependent image coordinates, so the same grid location across views generally corresponds to different scene content. We therefore use the current Gaussian scene geometry to establish cross-view correspondences between residual histories before constructing a shared reference.

Geometry-Constrained Correspondence. For source view v, let $\mathbf{p}_{v,a}$ denote the image-space center of residual-grid location a. Since the residual history contains no geometry, we derive a depth statistic $z_{v,a}^t$ from the depth-ordered Gaussian contributions of the current rendering and lift $\mathbf{p}_{v,a}$ to 3D. The resulting point is then reprojected into another training view $u \neq v$:

$$\mathbf{x}_{v,a}^t = \mathrm{BackProj}\left(\mathbf{p}_{v,a}, z_{v,a}^t; \pi_v\right), \qquad \left(\mathbf{q}_{u\leftarrow v,a}^t, \tilde{z}_{u\leftarrow v,a}^t\right) = \mathrm{Proj}\left(\mathbf{x}_{v,a}^t; \pi_u\right). \tag{4}$$

Here, BackProj and Proj denote calibrated camera back-projection and projection. We consider only target views with initialized residual histories and exclude the source view. A correspondence

is retained when the reprojected location lies inside the target image and its projected depth is consistent with the rendering-derived depth at that location. For each valid correspondence, we map $\mathbf{q}^t_{u\leftarrow v,a}$ to the residual-history grid and bilinearly sample $\mathbf{H}^t_u$ to obtain $\mathbf{h}^t_{u\leftarrow v,a}$.

Cross-View Residual Consensus. Let $\mathcal{V}^t_{v,a}$ denote the set of comparison views that satisfy the geometric consistency test for source cell a at iteration t. When $|\mathcal{V}^t_{v,a}| \geq 2$, we define the cross-view residual consensus as

$$\mathbf{C}^t_{v,a} = \mathrm{Med}\left\{\mathbf{h}^t_{u\leftarrow v,a} \mid u \in \mathcal{V}^t_{v,a}\right\}, \tag{5}$$

where Med denotes the component-wise median across valid comparison views. Cells with fewer than two valid comparison views receive no subsequent disagreement compensation. The resulting $\mathbf{C}^t_{v,a}$ summarizes agreement among geometrically corresponding residual histories and serves as a relative cross-view reference for $\mathbf{H}^t_{v,a}$. Importantly, $\mathbf{C}^t_{v,a}$ is a geometry-conditioned relative reference rather than an estimate of static error.

3.4 Consensus-Relative Residual Compensation

Consensus-Relative Disagreement. At iteration t, given the per-view residual history $\mathbf{H}^t_{v,a}$ and cross-view consensus reference $\mathbf{C}^t_{v,a}$, we define the consensus-relative disagreement at each grid location with valid consensus as

$$\mathbf{D}^t_{v,a} = \mathbf{H}^t_{v,a} - \mathbf{C}^t_{v,a}. \tag{6}$$

If no valid consensus is available, we set $\mathbf{D}^t_{v,a} = \mathbf{0}$, yielding no compensation at that location. Thus, $\mathbf{D}^t_{v,a}$ measures the departure of the source-view history from its geometry-conditioned cross-view reference. It is an operational residual-space quantity rather than a semantic transient indicator, and may reflect residual distractors, view-dependent appearance, modeling mismatch, or other view-specific discrepancies.

Compensated Rendering and Residual Reshaping. Collecting the disagreement over the residual grid gives $\mathbf{D}^t_v \in \mathbb{R}^{3\times K\times K}$. We bilinearly upsample it to image resolution using $\mathcal{U}(\cdot)$ and apply it as a detached additive correction:

$$\widetilde{\mathbf{I}}^t_v = \widehat{\mathbf{I}}^t_v + \mathrm{sg}\left(\mathcal{U}(\mathbf{D}^t_v)\right), \tag{7}$$

where $\mathrm{sg}(\cdot)$ leaves the correction unchanged in the forward pass while blocking gradients through it. The resulting photometric residual is

$$\widetilde{\mathbf{e}}^t_v = \mathbf{I}_v - \widetilde{\mathbf{I}}^t_v = \mathbf{e}^t_v - \mathrm{sg}\left(\mathcal{U}(\mathbf{D}^t_v)\right). \tag{8}$$

The native reconstruction loss is then evaluated using $\widetilde{\mathbf{I}}^t_v$ instead of $\widehat{\mathbf{I}}^t_v$, while the supervision weights $\boldsymbol{\omega}^t_v$, loss formulation, and other optimization terms remain unchanged. Although gradients do not propagate through the disagreement correction, they still flow through $\widehat{\mathbf{I}}^t_v$ to the Gaussian parameters. CoReSplat therefore reshapes the photometric residual before gradient formation rather than modifying gradients after they are formed.

This residual form clarifies the effect of consensus-relative compensation. At locations with valid consensus,

$$\widetilde{\mathbf{e}}^t_v = \mathbf{e}^t_v - \mathcal{U}\left(\mathbf{H}^t_v - \mathbf{C}^t_v\right)$$

in the forward pass. When $\mathbf{e}^t_v \approx \mathcal{U}(\mathbf{H}^t_v)$, the effective residual approaches the cross-view reference, $\widetilde{\mathbf{e}}^t_v \approx \mathcal{U}(\mathbf{C}^t_v)$; otherwise, the current-to-history deviation remains. Thus, CoReSplat compensates the historical departure from cross-view consensus without forcing the current residual to equal the consensus.

3.5 State Update and Inference

CoReSplat follows a temporally ordered state update. At iteration t, the history bank $\{\mathbf{H}^t_u\}_{u=1}^V$ is used to construct $\mathbf{C}^t_v$ and $\mathbf{D}^t_v$ for the sampled view v, producing the compensated rendering $\widetilde{\mathbf{I}}^t_v$. The reconstruction objective then updates the Gaussian scene from θ^t to θ^{t+1}. After the update, we re-render view v to obtain $\bar{\mathbf{e}}^{t+1}_v$ and accumulate it with the supervision weights $\boldsymbol{\omega}^t_v$ to update $\mathbf{H}^{t+1}_v$ as in Sec. 3.2. Thus, newly generated residual evidence affects only subsequent iterations, not the update that produced it.

All residual histories, correspondences, consensus references, and disagreement quantities are training-only. At inference, they are discarded, and novel views are rendered directly from the optimized Gaussian scene using the original 3DGS renderer.

4 EXPERIMENTS

4.1 EXPERIMENTAL SETUP

Datasets, Baselines, and Evaluation Metrics. We evaluate CoReSplat on *NeRF On-the-go* (Ren et al., 2024) and *RobustNeRF* (Sabour et al., 2023). For *NeRF On-the-go*, we use six standard scenes with the camera-undistortion protocol of RobustSplat (Fu et al., 2025b), downsampling *Patio* by $4\times$ and the others by $8\times$; all five *RobustNeRF* scenes are downsampled by $8\times$. Training uses distractor-containing images and evaluation uses clean test views. We compare with 3DGS (Kerbl et al., 2023), WildGaussians (Kulhanek et al., 2024), DeSplat (Wang et al., 2025b), SpotLessSplats (Sabour et al., 2025), DeGauss (Wang et al., 2025a), RobustSplat (Fu et al., 2025b), DualSplat (Wang et al., 2026), and 3DGS-HPC (Chen et al., 2026), using their official implementations and method-specific training settings under a common preprocessing and evaluation protocol. We report PSNR↑, SSIM↑, and LPIPS↓ on clean test views.

Implementation Details. For the primary configuration, we optimize the underlying reconstruction pipeline for 20k iterations using its native procedure, activate CoReSplat, and continue to the original stopping iteration. Residual histories use a 16×16 grid, with cross-view consensus requiring at least two geometrically valid comparison views and a relative depth-consistency threshold of 0.02. We further integrate CoReSplat into DeSplat (Wang et al., 2025b), RobustSplat (Fu et al., 2025b), and DualSplat (Wang et al., 2026), while retaining each method's native scene representation, renderer, supervision, loss, and optimization schedule.

4.2 QUANTITATIVE EVALUATION

Table 1 shows that CoReSplat integrated with DualSplat achieves the strongest overall performance on *NeRF On-the-go*. The resulting model obtains the highest PSNR on five of the six scenes and the lowest LPIPS on all six scenes. Relative to the native DualSplat model, CoReSplat further improves the mean PSNR by 0.337 dB while simultaneously improving the mean SSIM and LPIPS. These results indicate that residual-level compensation remains beneficial even after a strong distractor-aware reconstruction has already been established, and that the gain persists across scenes under different occlusion conditions.

RobustNeRF provides complementary evidence under a different benchmark setting. As reported in Table 2, RobustSplat with CoReSplat achieves the highest mean PSNR and the best PSNR on four of the five evaluated scenes, whereas DualSplat with CoReSplat attains the best mean SSIM and the lowest mean LPIPS. In particular, the perceptual improvements of the DualSplat integration are observed across all five scenes. Although the SSIM response varies more across the underlying reconstruction methods than on *NeRF On-the-go*, the PSNR gains of integrations remain positive across all evaluated method and scenes. The persistence of these gains on a second benchmark suggests that the benefit of CoReSplat is not specific to the occlusion-stratified *NeRF On-the-go* setting.

4.3 QUALITATIVE EVALUATION

Figure 2 further illustrates the differences in novel-view synthesis quality. On the *Corner* scene, several competing methods retain visible ghosting, haze, or local distortions in the highlighted floor and object regions, whereas the CoReSplat-enhanced reconstruction presents cleaner local appearance and sharper boundaries. A similar pattern can be observed on *Crab2*, where residual artifacts around object boundaries and previously disturbed regions are reduced while local scene structure is preserved.

Table 1: Quantitative comparison on *NeRF On-the-go*. We report PSNR↑, SSIM↑, and LPIPS↓ on six scenes spanning low, medium, and high distractor occlusion levels. CoReSplat integrations are shown directly below their respective underlying methods. The best unrounded result in each column is highlighted in bold.

Method	Low Occlusion						Medium Occlusion						High Occlusion						Mean		
	Mountain			Fountain			Corner			Patio			Spot			Patio-High					
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
3DGS	20.83	0.648	0.257	19.97	0.618	0.252	22.83	0.794	0.194	17.16	0.667	0.237	18.09	0.632	0.485	16.37	0.507	0.500	19.208	0.644	0.321
WildGaussians	19.63	0.681	0.253	20.48	0.661	0.224	25.35	0.862	0.111	21.07	0.804	0.134	24.87	0.872	0.099	22.60	0.801	0.170	22.333	0.780	0.165
SpotLessSplats	21.17	0.665	0.244	20.58	0.631	0.226	25.30	0.848	0.122	21.41	0.799	0.118	24.46	0.824	0.149	21.95	0.742	0.230	22.478	0.752	0.182
3DGS-HPC	21.26	0.734	0.152	21.45	0.703	0.143	24.81	0.886	0.075	20.46	0.805	0.112	23.95	0.893	0.083	22.68	0.831	0.125	22.435	0.809	0.115
DeGauss	21.25	0.741	0.177	18.80	0.651	0.204	26.39	0.891	0.068	21.78	0.822	0.098	26.58	0.906	0.062	22.97	0.823	0.126	22.962	0.806	0.123
DeSplat	20.25	0.709	0.158	20.40	0.672	0.165	26.16	0.882	0.076	18.71	0.775	0.138	25.58	0.890	0.096	22.48	0.823	0.137	22.263	0.792	0.128
+ CoReSplat	20.98	0.716	0.151	20.43	0.674	0.162	26.29	0.883	0.078	18.97	0.777	0.137	25.87	0.891	0.093	22.51	0.818	0.139	22.508	0.793	0.127
RobustSplat	21.31	0.731	0.204	20.89	0.701	0.199	26.34	0.898	0.102	21.44	0.827	0.140	25.31	0.906	0.102	22.65	0.836	0.146	22.990	0.817	0.149
+ CoReSplat	21.41	0.734	0.206	21.05	0.703	0.197	26.76	0.900	0.101	21.79	0.830	0.138	**26.65**	0.914	0.095	23.01	0.838	0.142	23.445	0.820	0.147
DualSplat	21.65	**0.765**	0.123	21.72	**0.721**	0.120	26.73	0.907	0.047	21.79	0.837	0.073	25.59	0.911	0.045	23.59	0.852	0.084	23.512	0.832	0.082
+ CoReSplat	**21.95**	0.758	**0.122**	**21.76**	**0.721**	**0.117**	**27.15**	**0.911**	**0.043**	**22.04**	**0.872**	**0.041**	26.56	**0.919**	**0.038**	**23.63**	**0.854**	**0.080**	**23.848**	**0.839**	**0.074**

Table 2: Quantitative comparison on *RobustNeRF*. We report PSNR↑, SSIM↑, and LPIPS↓ across five scenes. CoReSplat integrations are shown directly below their underlying methods. The best result in each column is highlighted in bold.

Method	Android			Crab1			Crab2			Statue			Yoda			Mean		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
3DGS	23.02	0.771	0.126	31.00	0.949	0.078	31.19	0.913	0.134	21.12	0.796	0.158	30.44	0.917	0.138	27.354	0.869	0.127
WildGaussians	24.76	0.826	0.090	31.14	0.950	0.076	30.51	0.904	0.153	22.52	0.859	0.117	30.85	0.911	0.148	27.956	0.890	0.117
SpotLessSplats	24.27	0.809	0.088	34.93	0.964	0.060	34.24	0.931	0.093	22.26	0.817	0.118	34.64	0.937	0.106	30.068	0.892	0.093
3DGS-HPC	24.52	0.825	0.069	33.90	0.968	0.050	34.44	0.938	0.095	22.77	0.862	0.090	35.08	0.943	0.089	30.142	0.907	0.079
DeGauss	24.82	0.830	0.120	35.96	0.968	0.097	34.82	0.937	0.153	22.75	0.864	0.108	34.72	0.941	0.149	30.614	0.908	0.125
DeSplat	24.24	0.819	0.089	36.09	0.969	0.045	34.13	0.931	0.098	22.65	0.850	0.096	34.03	0.934	0.099	30.228	0.901	0.085
+ CoReSplat	24.28	0.817	0.090	36.24	0.969	0.040	34.32	0.929	0.099	22.82	0.850	0.095	34.18	0.935	0.091	30.368	0.900	0.083
RobustSplat	24.69	0.831	0.124	36.51	0.971	0.096	34.72	**0.940**	0.153	22.70	0.865	0.108	35.15	0.944	0.152	30.754	0.910	0.127
+ CoReSplat	**24.84**	**0.834**	0.122	**37.14**	0.972	0.094	**34.84**	**0.940**	0.153	22.82	**0.869**	0.105	**35.26**	**0.945**	0.151	**30.980**	0.912	0.125
DualSplat	24.64	0.830	0.065	36.68	0.972	0.055	34.55	**0.940**	0.091	22.88	0.868	0.065	34.81	0.944	0.090	30.712	0.911	0.073
+ CoReSplat	24.79	0.831	**0.050**	37.10	**0.990**	**0.002**	34.63	**0.940**	**0.031**	**22.96**	0.868	**0.062**	35.09	0.944	**0.029**	30.914	**0.915**	**0.035**

4.4 Cross-Baseline Integration

To evaluate whether the proposed compensation mechanism provides consistent benefits beyond a specific reconstruction pipeline, we compare each underlying method before and after integrating CoReSplat. Table 3 reveals a consistent improvement pattern. CoReSplat achieves positive PSNR gains on all evaluated method-scene combinations across two benchmarks. These improvements across different scenes, occlusion conditions, and datasets indicate that the benefit of consensus-relative residual compensation is not restricted to a specific reconstruction setting. Figure 3 provides visual evidence consistent with the quantitative results. Across DeSplat, DualSplat, and RobustSplat, integrating CoReSplat reduces visible residual artifacts in the highlighted regions while maintaining the overall scene structures reconstructed by the underlying methods.

4.5 Ablation Study

We conduct ablation studies on all six *NeRF On-the-go* scenes using RobustSplat as the underlying reconstruction baseline. All variants are initialized from the same RobustSplat model and evaluated under the same training protocol, enabling a controlled analysis of the main design choices of CoReSplat. The results are summarized in Table 4.

Effect of the Consensus-Relative Design. To examine whether accumulated residual history alone is sufficient, we replace the consensus-relative disagreement $D = H - C$ with direct-history compensation $D = H$. As shown in Table 4, this variant consistently degrades reconstruction quality, supporting the use of cross-view consensus as a relative reference rather than compensating the entire residual history directly. This comparison validates the core consensus-relative design of CoReSplat, where residual evidence is interpreted with respect to cross-view agreement before compensation.

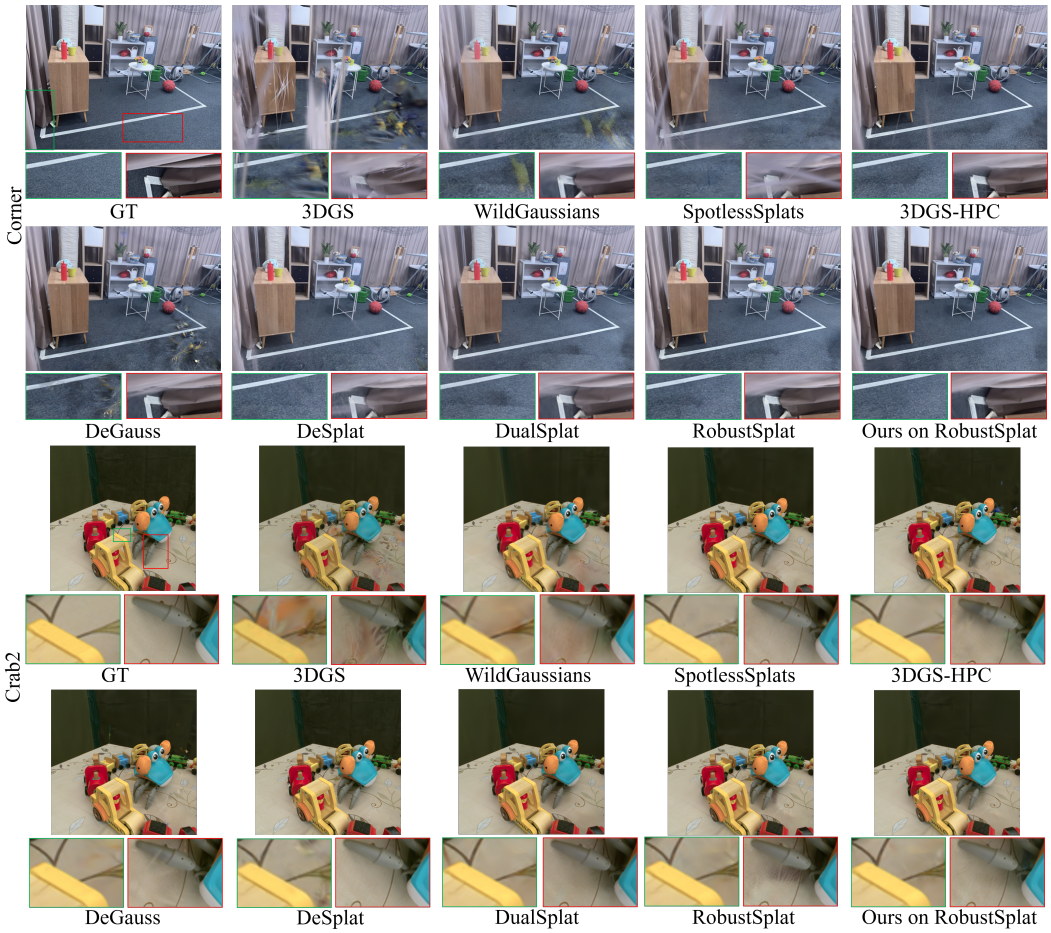

Figure 2: Qualitative comparison with existing methods. We show novel-view synthesis results on *Corner* from *NeRF On-the-go* and *Crab2* from *RobustNeRF*. CoReSplat is visualized with RobustSplat as a representative underlying integration. The highlighted regions reveal differences in residual artifacts and local structure preservation, where CoReSplat produces cleaner local appearance and sharper scene details than the original underlying method and competing baselines.

Table 3: Cross-baseline PSNR gains from CoReSplat integration. We report scene-wise PSNR improvements (ΔPSNR) obtained by integrating CoReSplat into DeSplat, RobustSplat, and DualSplat on *NeRF On-the-go* and *RobustNeRF*. All evaluated method pairs exhibit positive gains, and the largest gain in each column is highlighted in bold.

Method	*NeRF On-the-go*							*RobustNeRF*					
	Mountain	Fountain	Corner	Patio	Spot	Patio-High	Mean	Android	Crab1	Crab2	Statue	Yoda	Mean
RobustSplat + CoReSplat	+0.097	**+0.158**	**+0.422**	**+0.346**	**+1.343**	**+0.360**	**+0.454**	+0.146	**+0.623**	+0.119	+0.129	+0.112	**+0.226**
DeSplat + CoReSplat	**+0.730**	+0.035	+0.131	+0.261	+0.290	+0.020	+0.245	+0.039	+0.146	**+0.184**	**+0.165**	+0.149	+0.137
DualSplat + CoReSplat	+0.300	+0.040	+0.420	+0.251	+0.968	+0.041	+0.337	**+0.154**	+0.425	+0.083	+0.083	**+0.280**	+0.205

Effect of Cross-Iteration Residual Evidence. We next replace the accumulated residual history with only the latest residual observation. Table 4 shows that this variant yields lower PSNR and SSIM but slightly better LPIPS. Overall, the results support cross-iteration accumulation as a more effective source of residual evidence for maintaining reconstruction fidelity and structural similarity than relying on a single optimization state.

Importance of Supervision Alignment. Finally, we remove the native supervision weights when accumulating residual history. As shown in Table 4, this variant degrades reconstruction quality across all three metrics, indicating that residual evidence is more informative when its accumulation

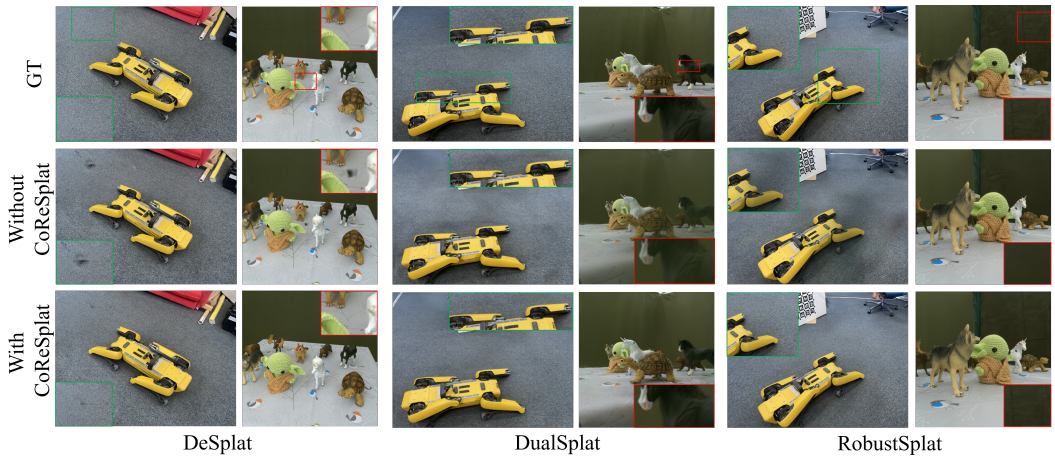

Figure 3: Qualitative effect of cross-baseline CoReSplat integration. We compare native reconstructions and their CoReSplat-enhanced counterparts for DeSplat, DualSplat, and RobustSplat, with the corresponding ground truth shown above. In each column, the left result corresponds to *Spot*, and the right example to *Yoda*. We provide highlighted regions to illustrate reduced residual artifacts and cleaner local structures after integrating CoReSplat across different underlying reconstruction pipelines.

Table 4: Ablation and design analysis of CoReSplat on *NeRF On-the-go*. Results are averaged over all six scenes, with RobustSplat as the underlying reconstruction method.

Variant	PSNR↑	SSIM↑	LPIPS↓
Native RobustSplat	22.990	0.817	0.149
Direct-history compensation	23.144	0.816	0.148
Latest-only residual evidence	23.241	0.817	**0.147**
No supervision alignment	23.216	0.817	0.149
CoReSplat	**23.445**	**0.820**	**0.147**

follows the same supervision support that drives the underlying reconstruction. This alignment prevents downweighted observations from contributing disproportionately to the historical evidence used for subsequent consensus construction.

Overall, these ablations validate the main design choices of CoReSplat, including cross-iteration residual evidence, consensus-relative interpretation, and alignment with the native supervision used for scene optimization.

5 Conclusion

Distractor-free 3D reconstruction requires more than determining which observations should participate in optimization: even retained observations may produce photometric discrepancies whose influence on the shared scene remains ambiguous. CoReSplat addresses this supervision-to-update ambiguity by combining cross-iteration residual evidence with cross-view consensus to establish a relative reference, from which consensus-relative disagreement is derived and compensated before Gaussian optimization. This formulation recalibrates the photometric discrepancy presented to scene optimization, without treating the consensus as static ground truth or the disagreement as a transient semantic label. Experiments on NeRF On-the-go and RobustNeRF, together with integrations into multiple distractor-free 3DGS pipelines, demonstrate improved novel-view synthesis across different reconstruction strategies. More broadly, our findings highlight that robust reconstruction depends not only on identifying reliable supervision, but also on determining how the residuals of retained observations should shape the shared scene representation.

AI USE STATEMENT

In this work, generative AI tools were used to support manuscript editing and refinement, as well as to facilitate the exploration and identification of relevant literature. The retrieved references and related-work suggestions assisted by AI tools were subsequently verified by the authors through manual examination of the corresponding publications. Generative AI tools were not involved in producing experimental results, executing experiments, developing the proposed approach, or formulating scientific conclusions. All AI-assisted modifications to the manuscript were carefully reviewed by the authors to ensure accuracy and alignment with the actual contributions of this work. The authors assume full responsibility for the final content of this submission.

REFERENCES

Jonathan T Barron, Ben Mildenhall, Matthew Tancik, Peter Hedman, Ricardo Martin-Brualla, and Pratul P Srinivasan. Mip-NeRF: A multiscale representation for anti-aliasing neural radiance fields. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 5835–5844. IEEE, 2021.

Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. Mip-NeRF 360: Unbounded anti-aliased neural radiance fields. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5460–5469. IEEE, 2022.

Jiahao Chen, Yipeng Qin, Lingjie Liu, Jiangbo Lu, and Guanbin Li. NeRF-HuGS: Improved neural radiance fields in non-static scenes using heuristics-guided segmentation. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 19436–19446. IEEE, 2024.

Jiahao Chen, Yipeng Qin, Ganlong Zhao, Xin Li, Wenping Wang, and Guanbin Li. 3DGS-HPC: Distractor-free 3D gaussian splatting with hybrid patch-wise classification. In *Proceedings of the 43rd International Conference on Machine Learning (ICML)*, pp. 1–19, 2026.

Xingyu Chen, Qi Zhang, Xiaoyu Li, Yue Chen, Ying Feng, Xuan Wang, and Jue Wang. Hallucinated neural radiance fields in the wild. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12933–12942. IEEE, 2022.

Hiba Dahmani, Moussab Bennehar, Nathan Piasco, Luis Roldao, and Dzmitry Tsishkou. SWAG: Splatting in the wild images with appearance-conditioned gaussians. In *European Conference on Computer Vision*, pp. 325–340. Springer, 2024.

Xiaobiao Du, Yida Wang, and Xin Yu. MVGS: Multi-view regulated gaussian splatting for novel view synthesis. In *European Conference on Computer Vision*, pp. 564–581. Springer, 2026.

Chuanyu Fu, Guanying Chen, Yuqi Zhang, Kunbin Yao, Yuan Xiong, Chuan Huang, Shuguang Cui, Yasuyuki Matsushita, and Xiaochun Cao. RobustSplat++: Decoupling densification, dynamics, and illumination for in-the-wild 3DGS. *arXiv preprint arXiv:2512.04815*, 2025a.

Chuanyu Fu, Yuqi Zhang, Kunbin Yao, Guanying Chen, Yuan Xiong, Chuan Huang, Shuguang Cui, and Xiaochun Cao. RobustSplat: Decoupling densification and dynamics for transient-free 3DGS. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 27126–27136. IEEE, 2025b.

Yuto Furutani, Takashi Otonari, Kaede Shiohara, and Toshihiko Yamasaki. Wild3R: Feed-forward 3D gaussian splatting from unconstrained sparse photo collection. *arXiv preprint arXiv:2606.11894*, 2026.

Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2D gaussian splatting for geometrically accurate radiance fields. In *ACM SIGGRAPH 2024 conference papers*, pp. 1–11, 2024.

Zeji Hui, Amirali Khodadadian Gostar, WeiQin Chuah, Alireza Bab-Hadiashar, and Ruwan Tennakoon. Joint modeling of corruption-driven and information-limited uncertainty for robust 3D gaussian splatting. In *2026 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 688–697. IEEE, 2026.

Yulei Kang, Tianze Zhu, Jian-Fang Hu, Jianhuang Lai, and Wei-Shi Zheng. HarmoGS: Robust 3D gaussian splatting in the wild via conflict-aware gradient harmonization. *arXiv preprint arXiv:2605.13073*, 2026.

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3D gaussian splatting for real-time radiance field rendering. *ACM Transation on Graphics*, 42(4):139–1, 2023.

Hanyang Kong, Xingyi Yang, and Xinchao Wang. RoGSplat: Robust gaussian splatting via generative priors. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 25735–25745. IEEE, 2025.

Jonas Kulhanek, Songyou Peng, Zuzana Kukelova, Marc Pollefeys, and Torsten Sattler. WildGaussians: 3D gaussian splatting in the wild. In *Proceedings of the 38th International Conference on Neural Information Processing Systems*, volume 37, pp. 21271–21288, 2024.

Chengqi Li, Zhihao Shi, Yangdi Lu, Wenbo He, and Xiangyu Xu. Robust neural rendering in the wild with asymmetric dual 3D gaussian splatting. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, volume 38, Main Conference, pp. 50166–50191, 2025.

Peihao Li, Shaohui Wang, Chen Yang, Bingbing Liu, Weichao Qiu, and Haoqian Wang. NeRF-MS: Neural radiance fields with multi-sequence. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 18545–18554. IEEE, 2023.

Ruiqi Li and Yiu-ming Cheung. Modeling and identifying distractors with curriculum for robust 3D gaussian splatting. In *Proceedings of the 33rd ACM International Conference on Multimedia*, pp. 10122–10131, 2025.

Shengjia Liang, Tong Chen, Qichuan Geng, Yuan Xiong, and Zhong Zhou. CleanSplat: curriculum structural gaussian splatting for spot-free novel view synthesis. *Virtual Reality & Intelligent Hardware*, 8(2):235–249, 2026.

Jingyu Lin, Jiaqi Gu, Lubin Fan, Bojian Wu, Yujing Lou, Renjie Chen, Ligang Liu, and Jieping Ye. HybridGS: Decoupling transients and statics with 2D and 3D gaussian splatting. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 788–797. IEEE, 2025.

Tao Lu, Mulin Yu, Linning Xu, Yuanbo Xiangli, Limin Wang, Dahua Lin, and Bo Dai. Scaffold-GS: Structured 3D gaussians for view-adaptive rendering. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20654–20664. IEEE, 2024.

Alexander Markin, Vadim Pryadilshchikov, Artem Komarichev, Ruslan Rakhimov, Peter Wonka, and Evgeny Burnaev. T-3DGS: Removing transient objects for 3D scene reconstruction. *arXiv preprint arXiv:2412.00155*, 2024.

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 *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7206–7215, 2021.

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. In *European Conference Computer Vision*, pp. 405–421, 2020.

Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics*, 41(4):1–15, 2022.

Tianbo Pan, Xingyi Yang, Shizun Wang, and Xinchao Wang. Visual geometry transformer in the wild: Distractor-free 3D reconstruction. *arXiv preprint arXiv:2606.22787*, 2026.

Wongi Park, Jiyeon Lim, Minjae Lee, Myeongseok Nam, Seongjun Choi, Jungwoo Kim, Soomok Lee, William J Beksi, and Sang-Hyun Lee. Rectifying mask via entropy for distractor-free 3DGS in ambiguous scenarios. *arXiv preprint arXiv:2606.29496*, 2026a.

Wongi Park, Myeongseok Nam, Siwon Kim, Sangwoo Jo, and Soomok Lee. ForestSplats: Deformable transient field for gaussian splatting in the wild. In *2026 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 6978–6987. IEEE, 2026b.

Weining Ren, Zihan Zhu, Boyang Sun, Jiaqi Chen, Marc Pollefeys, and Songyou Peng. NeRF On-the-go: Exploiting uncertainty for distractor-free nerfs in the wild. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8931–8940. IEEE, 2024.

Sara Sabour, Suhani Vora, Daniel Duckworth, Ivan Krasin, David J Fleet, and Andrea Tagliasacchi. RobustNeRF: Ignoring distractors with robust losses. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20626–20636. IEEE, 2023.

Sara Sabour, Lily Goli, George Kopanas, Mark Matthews, Dmitry Lagun, Leonidas Guibas, Alec Jacobson, David Fleet, and Andrea Tagliasacchi. SpotLessSplats: Ignoring distractors in 3D gaussian splatting. *ACM Transactions on Graphics*, 44(2):1–11, 2025.

Kangmin Seo, MinKyu Lee, Tae-Young Kim, ByeongCheol Lee, JoonSeoung An, and Jae-Pil Heo. PDF-GS: Progressive distractor filtering for robust 3D gaussian splatting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Findings*, pp. 468–477, June 2026.

Jiadong Tang, Yu Gao, Dianyi Yang, Liqi Yan, Yufeng Yue, and Yi Yang. DroneSplat: 3D gaussian splatting for robust 3D reconstruction from in-the-wild drone imagery. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 833–843. IEEE, 2025.

Paul Ungermann, Armin Ettenhofer, Matthias Nießner, and Barbara Roessle. Robust 3D gaussian splatting for novel view synthesis in presence of distractors. In *DAGM German Conference on Pattern Recognition*, pp. 153–167. Springer, 2024.

Rui Wang, Quentin Lohmeyer, Mirko Meboldt, and Siyu Tang. DeGauss: Dynamic-static decomposition with gaussian splatting for distractor-free 3D reconstruction. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 6294–6303. IEEE, 2025a.

Xu Wang, Zhiru Wang, Shiyun Xie, Chengwei Pan, and Yisong Chen. DualSplat: Robust 3D gaussian splatting via pseudo-mask bootstrapping from reconstruction failures. *arXiv preprint arXiv:2604.21631*, 2026.

Yihao Wang, Marcus Klasson, Matias Turkulainen, Shuzhe Wang, Juho Kannala, and Arno Solin. DeSplat: Decomposed gaussian splatting for distractor-free rendering. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 722–732. IEEE, 2025b.

Yukun Wang, Kunhong Li, Minglin Chen, Longguang Wang, Shunbo Zhou, Kaiwen Xue, and Yulan Guo. Distractor-free novel view synthesis via exploiting memorization effect in optimization. In *European Conference on Computer Vision*, pp. 477–493. Springer, 2024a.

Yuze Wang, Junyi Wang, and Yue Qi. WE-GS: An in-the-wild efficient 3D gaussian representation for unconstrained photo collections. *arXiv preprint arXiv:2406.02407*, 2024b.

Tianhao Wu, Fangcheng Zhong, Andrea Tagliasacchi, Forrester Cole, and Cengiz Oztireli. D^2NeRF: Self-supervised decoupling of dynamic and static objects from a monocular video. *Advances in Neural Information Processing Systems*, 35:32653–32666, 2022.

Congrong Xu, Justin Kerr, and Angjoo Kanazawa. Splatfacto-W: A nerfstudio implementation of gaussian splatting for unconstrained photo collections. *arXiv preprint arXiv:2407.12306*, 2024.

Yifan Yang, Shuhai Zhang, Zixiong Huang, Yubing Zhang, and Mingkui Tan. Cross-ray neural radiance fields for novel-view synthesis from unconstrained image collections. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 15855–15865. IEEE, 2023.

Yunseo Yang, Youngho Yoon, and Kuk-Jin Yoon. PriSplat: Propagating reliable multi-view information for distractor-free 3DGS. In *European Conference on Computer Vision*, 2026.

Zehao Yu, Anpei Chen, Binbin Huang, Torsten Sattler, and Andreas Geiger. Mip-Splatting: Alias-free 3D gaussian splatting. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 19447–19456. IEEE, 2024.

Dongbin Zhang, Chuming Wang, Weitao Wang, Peihao Li, Minghan Qin, and Haoqian Wang. Gaussian in the wild: 3D gaussian splatting for unconstrained image collections. In *European Conference on Computer Vision*, pp. 341–359. Springer, 2024.