Where Fixes Really Help: Cross-Fitted Repair-Model Supervision for 3DGS

Anonymous authors
Paper under double-blind review

Abstract

Sparse-view 3D Gaussian Splatting (3DGS) is severely underconstrained away from observed cameras, often producing unstable geometry and appearance. Pretrained repair models can enrich supervision by correcting rendered novel views, yet their outputs may also introduce scene-inconsistent content that 3DGS absorbs into persistent artifacts. The resulting challenge is to determine which repaired pixels provide reliable supervision. We introduce FixTrust-GS, a framework that estimates pixel-wise repair reliability from the captured scene. FixTrust-GS constructs cross-fitted auxiliary reconstructions using only captured images and evaluates whether each repaired pixel agrees with their predictions better than the corresponding initial render. This relative agreement improvement yields dense weights for continued 3DGS optimization, retaining supported repair content while suppressing inconsistent edits. The framework requires only rendered–repaired image pairs and works with pretrained repair models without modifying their parameters. Experiments across three datasets, two distinct repair models, and three sparse-view budgets show that FixTrust-GS improves reconstruction fidelity over uniform whole-image weighting while often using fewer Gaussians in the final model. Code: `https://anonymous.4open.science/r/FixTrust-GS-1D4B`.

1 Introduction

Reconstructing a photorealistic 3D scene from images is one of the central tasks in computer vision and graphics, tested with cameras absent from the capture by novel view synthesis (NVS). Neural Radiance Fields (NeRFs) established differentiable scene reconstruction through continuous radiance fields and volume rendering (Mildenhall et al., 2020), and Mip-NeRF 360 extended this formulation to challenging unbounded scenes (Barron et al., 2022). More recently, 3D Gaussian Splatting (3DGS) has become a leading NVS representation by optimizing explicit anisotropic Gaussians and rendering them with a visibility-aware rasterizer (Kerbl et al., 2023). Follow-up work improves anti-aliasing and geometric accuracy while retaining efficient rasterization (Yu et al., 2024; Huang et al., 2024). Its combination of rendering quality, fast optimization, and real-time display makes 3DGS a practical foundation for scene reconstruction.

Sparse capture, however, makes this inverse problem severely underconstrained: with only a few cameras, many 3D explanations reproduce the training images but disagree in unseen regions, appearing as floaters, misplaced surfaces, blurred texture, and collapsed backgrounds. One family of methods adds regularization or priors, from unseen-view constraints and depth supervision to agreement between independently trained models (Niemeyer et al., 2022; Yang et al., 2023; Li et al., 2024; Zhu et al., 2024; Xiong et al., 2025; Zhang et al., 2024; Xu et al., 2025b; Charatan et al., 2024; Chen et al., 2024; Xu et al., 2025a). These approaches improve geometry and appearance, yet weakly observed regions remain ambiguous.

Generative repair models offer a complementary source of supervision: the pipeline renders novel cameras from an initial reconstruction, repairs them with an image or video diffusion model, and uses the outputs as pseudo-views (Wynn & Turmukhambetov, 2023; Wu et al., 2024; Liu et al., 2024; Zhong et al., 2025; Wu et al., 2025; Wei et al., 2026). The repair prior supplies realistic textures and fills missing regions, but plausibility is not scene fidelity: it may shift an edge, alter an object,

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.

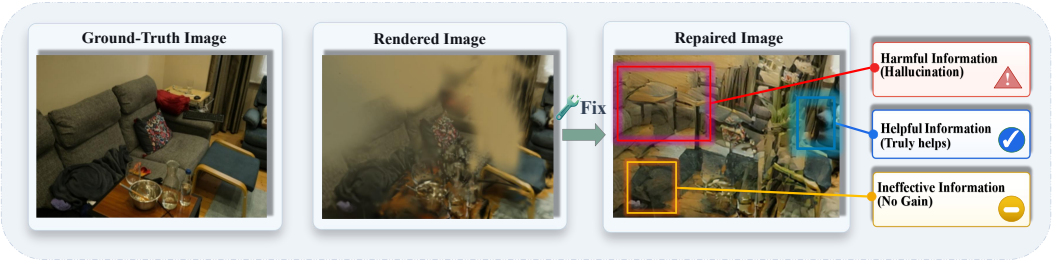

Figure 1: **Three kinds of information can coexist in one repaired view.** Based on our observations, a repair may supply helpful information (blue), harmful information that hallucinates scene content (red), and ineffective information that offers no gain over the initial render (orange). The ground-truth image is shown only to explain these categories; it is unavailable to FixTrust-GS.

or complete an incorrect surface, and repair supervision turns a local 2D hallucination into coherent but incorrect 3D geometry.

Existing strategies for integrating repaired views range from uniform supervision to confidence-based weighting. Confidence estimates draw on camera distance and Gaussian scale (Liu et al., 2024), Fisher information (Zhou et al., 2026), and expected information gain (Wang et al., 2026). HAD predicts hallucinations, while ConFixGS checks edits by reprojection (Liu et al., 2026; Song et al., 2026). A repaired view combines content inherited from the initial rendering with changes introduced by the generative prior. Based on our observations, we classify repaired content into three categories (Figure 1): **helpful** information that moves a region toward captured-scene evidence, **harmful** information that introduces scene-inconsistent hallucinations, and **ineffective** information that leaves the reconstruction no better supported. These categories can coexist within one repaired image, yet an image-level weight scales all three together. We therefore frame repair integration as pixel-wise acceptance: how much supervision should each repaired pixel receive?

The key is to assess an edit *relative to the render it changes*. An absolute residual can be large simply because a sparse-view reconstruction is uncertain, so its magnitude alone does not establish whether a repair improves on the initial render. We compare the pre- and post-repair pixels against the same set of scene-specific predictions at the same camera. Bias shared by those predictions affects both candidates, while a repair receives credit when it reduces their mean disagreement. This paired test provides observable relative evidence for repair quality: changes with positive average support receive supervision, others receive zero weight, and unchanged regions contribute no artificial gain.

In this paper, we introduce FixTrust-GS to approximate this test with complementary K-fold 3DGS models trained only on real photographs. At each repair camera, their predictions measure relative error reduction from the sealed to the repaired pixel; its positive part becomes the supervision weight. The method requires only aligned RGB inputs and outputs and works with distinct pretrained repair models—Difix3D and DiffusionHarmonizer—without modifying their parameters or inference procedures (Wu et al., 2025; Zhang et al., 2026).

Our contributions are summarized as follows:

- **Cross-fitted localization of reliable repairs.** We formulate repair trust as a scene-specific comparison against complementary 3DGS reconstructions trained on captured views, identifying where repaired pixels are better supported than the original render.
- **Pixel-wise reliability-weighted supervision.** We convert the positive agreement improvement into a bounded dense reliability map, applied before the spatial reduction of both L1 and D-SSIM objectives; the loss retains supported corrections, suppresses inconsistent content, and preserves full supervision from captured images.
- **Empirical evaluation.** Across repair models, datasets, and three sparse-view budgets, FixTrust-GS improves fidelity over whole-image weighting, often with fewer Gaussians, and exposes divergence between perceptual quality and scene fidelity.

2 Related Work

Sparse-view neural reconstruction. Sparse inputs leave novel views underconstrained. NeRF-based methods address this regime through unseen-view regularization, frequency control, and learned RGBD priors (Niemeyer et al., 2022; Yang et al., 2023; Wynn & Turmukhambetov, 2023). Sparse-view 3DGS methods instead combine explicit Gaussians with monocular depth, guided densification, co-regularization, unseen-view constraints, pruning, or random dropout (Li et al., 2024; Zhu et al., 2024; Zhang et al., 2024; Xiong et al., 2025; Xu et al., 2025b). Generalizable models learn multi-scene priors and predict Gaussian representations from image pairs or multi-view matching features (Charatan et al., 2024; Chen et al., 2024; Xu et al., 2025a).

Reliability-aware 3DGS. Reliability-aware 3DGS reduces the influence of inconsistent observations or unstable scene elements. Methods for uncontrolled captures use learned uncertainty, pretrained features, scene decomposition, or transient masks to suppress distractors (Kulhanek et al., 2024; Sabour et al., 2024; Wang et al., 2025; Fu et al., 2025). A complementary line estimates uncertainty in the representation through parameter distributions, second-order sensitivity, or cues from optimization, viewing coverage, and Gaussian geometry (Savant Aira et al., 2025; Hanson et al., 2025; Hui et al., 2026). These approaches identify unreliable observations, renderings, or primitives from model-internal and multi-view cues.

Repair models for neural reconstruction. Generative priors first served as optimization-time regularizers or repeated novel-view synthesis (Wynn & Turmukhambetov, 2023; Wu et al., 2024). Newer pipelines repair rendered novel views and use them as 3D supervision. Difix3D and GSFix3D emphasize repair generation and staged integration, while DiffusionHarmonizer performs online enhancement of reconstructed views (Wu et al., 2025; Wei et al., 2026; Zhang et al., 2026). 3DGS-Enhancer weights supervision using camera distance and rendered Gaussian scale (Liu et al., 2024); FreeFix derives a confidence mask from Fisher information (Zhou et al., 2026); and FaithFusion estimates pixel-wise information gain (Wang et al., 2026). HAD predicts hallucination scores (Liu et al., 2026). ConFixGS uses reprojection, OracleGS validates with multi-view stereo evidence, and RogSplat compares pretrained features (Song et al., 2026; Topaloğlu et al., 2026; Kong et al., 2025). Together, these methods move repair-based reconstruction toward confidence-aware integration.

3 Method

Uncertainty indicates where a reconstruction lacks evidence; repair reliability asks whether the content supplied there agrees with the scene. FixTrust-GS estimates the latter by comparing each repaired pixel with complementary reconstructions of the observed photographs. The resulting positive relative error reduction weights spatial L1 and D-SSIM during continued optimization. Figure 2 summarizes the three stages: cross-fitted 3DGS models provide real-image-constrained evidence, the initial and repaired pixels are compared against that evidence at the same camera, and the resulting reliability map weights continued training from the sealed main state. Algorithm 1 gives the corresponding procedure. Appendix A consolidates the notation used below, and Appendix B reviews the 3DGS representation, rendering equations, and standard photometric objective.

3.1 Problem formulation: repair supervision from captured-view evidence

Let $\mathcal{T} = \{(v_n, I_n)\}_{n=1}^{N_t}$ denote the observed training views, where v_n is a calibrated camera and $I_n \in [0,1]^{H\times W\times 3}$ is its RGB image. We write $\mathcal{R}(\theta; v)$ for the differentiable rendering of a 3DGS with parameters θ from camera v. Optimizing all observed views to a designated repair step yields the sealed reconstruction θ^0 and its complete optimization state.

Consider a set of N_r repair cameras $\mathcal{V}_r = \{\widetilde{v}_j\}_{j=1}^{N_r}$. At the sealed state, we render the initial novel view and apply a pretrained repair model whose parameters remain fixed:

$$S_j = \mathcal{R}(\theta^0; \widetilde{v}_j), \qquad F_j = \mathrm{Fix}_\phi(S_j; \xi_j), \qquad j = 1, \ldots, N_r. \tag{1}$$

Here ϕ is frozen and ξ_j denotes any model-specific conditioning, such as a nearby reference image or temporal context. The repaired RGB F_j is kept fixed thereafter. Ground-truth RGB at $\widetilde{v}_j$ is

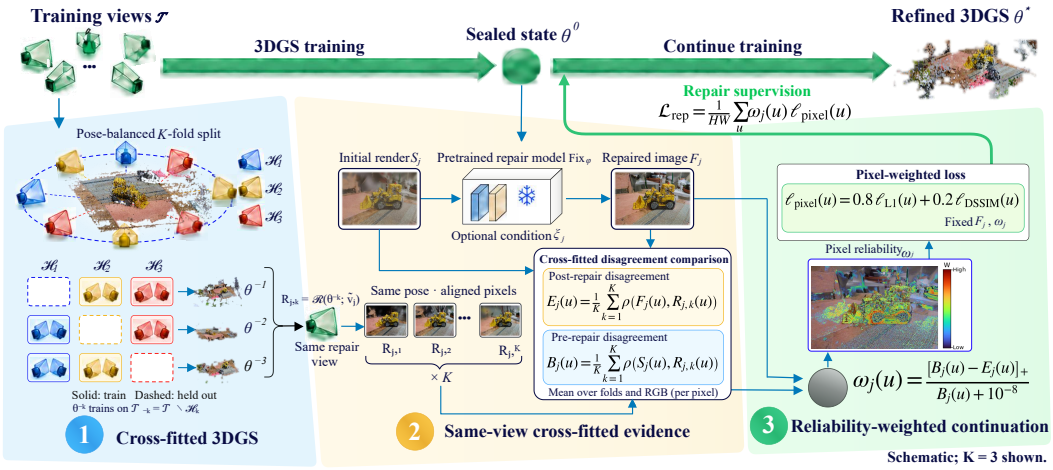

Figure 2: **Overview of FixTrust-GS.** (1) A pose-balanced split trains complementary 3DGS models using only captured photographs. (2) At each repair camera, the same cross-fitted renders assess the sealed render S_j and repaired image F_j, producing pre- and post-repair disagreements B_j and E_j. (3) Their positive relative reduction yields pixel reliability ω_j, which weights repair supervision while the sealed main 3DGS continues to $\theta^\star$. The diagram shows $K = 3$; the repair model, repaired images, and reliability maps remain fixed during continuation.

Algorithm 1 Cross-fitted repair-model supervision with FixTrust-GS

Require: Observed views $\mathcal{T}$; repair cameras $\mathcal{V}_r$; pretrained repair model Fix_ϕ; folds $K \geq 2$
Ensure: Refined 3DGS parameters $\theta^\star$

Stage I: Generate candidate repair supervision
$\theta^0 \leftarrow \textsc{Train3DGS}(\mathcal{T})$; seal the full training state
for all $\widetilde{v}_j \in \mathcal{V}_r$ **do**
  $S_j \leftarrow \mathcal{R}(\theta^0; \widetilde{v}_j)$; $F_j \leftarrow \text{Fix}_\phi(S_j; \xi_j)$
end for

Stage II: Build evidence using only observed photographs
Partition $\{1, \ldots, N_t\}$ into balanced index folds $\{\mathcal{H}_k\}_{k=1}^K$
for $k = 1$ to K **do**
  $\mathcal{T}_{-k} \leftarrow \{(v_n, I_n) : n \notin \mathcal{H}_k\}$; $\theta^{-k} \leftarrow \textsc{Train3DGS}(\mathcal{T}_{-k})$
  Render $R_{j,k} \leftarrow \mathcal{R}(\theta^{-k}; \widetilde{v}_j)$ for all $\widetilde{v}_j \in \mathcal{V}_r$
end for

Stage III: Evaluate edits and weight repair supervision
for all j and pixels u **do**
  $B_j(u) \leftarrow K^{-1}\sum_k \rho(S_j(u), R_{j,k}(u))$
  $E_j(u) \leftarrow K^{-1}\sum_k \rho(F_j(u), R_{j,k}(u))$
  $\omega_j(u) \leftarrow [B_j(u) - E_j(u)]_+ / (B_j(u) + \varepsilon)$
end for
Assign unit maps to real images and ω_j to repaired image F_j
$\theta^\star \leftarrow \textsc{Continue3DGS}(\theta^0, \mathcal{T}, \{(\widetilde{v}_j, F_j, \omega_j)\}_{j=1}^{N_r})$
return $\theta^\star$

unavailable to the method. Our objective is therefore to infer a dense map

$$\omega_j : \Omega \to [0, 1], \qquad \Omega = \{1, \ldots, H\} \times \{1, \ldots, W\}, \tag{2}$$

using the observed views, the initial render S_j, and the repair F_j. The weight should be large where the repair is better supported than the initial render and vanish where the edit is contradicted by the available reconstruction evidence.

Plug-and-play repair interface. The reliability construction treats Fix_ϕ as an inference-only mapping and is agnostic to its architecture, training data, and number of denoising steps. Its interface consists of F_j aligned with S_j at the same camera and resolution. The same cross-fitted weighting

and continuation stages therefore support Difix3D, DiffusionHarmonizer, and other pretrained repair models through aligned RGB outputs (Wu et al., 2025; Zhang et al., 2026).

3.2 Cross-fitted rendering consensus

A single reconstruction is unsuitable as its own reference: using S_j to judge F_j would reject an edit merely because it differs from the render it is intended to correct. We instead obtain references constrained only by captured photographs by cross-fitting the observed views. Let $\{\mathcal{H}_k\}_{k=1}^{K}$ be a disjoint, balanced partition of their indices:

$$\bigcup_{k=1}^{K} \mathcal{H}_k = \{1, \dots, N_t\}, \qquad \mathcal{H}_k \cap \mathcal{H}_{k'} = \varnothing \quad (k \neq k'). \tag{3}$$

For fold k, we remove $\mathcal{H}_k$ and train an auxiliary 3DGS on the complementary real-view subset,

$$\mathcal{T}_{-k} = \{(v_n, I_n) : n \notin \mathcal{H}_k\}, \qquad \theta^{-k} = \mathrm{Train3DGS}(\mathcal{T}_{-k}). \tag{4}$$

All auxiliary reconstructions use the same optimization recipe and terminate at the same repair step as the main reconstruction. When $N_t = K$, this construction reduces to leave-one-view-out training; for larger training sets, each fold holds out a balanced subset. We use $K = 3$ folds by default, while the formulation is independent of the number of observed views. In practice, we construct the folds deterministically from camera poses to balance their spatial coverage; the precise criterion is given in Appendix A.1.

Each auxiliary model is then rendered at every repair camera:

$$R_{j,k} = \mathcal{R}(\theta^{-k}; \widetilde{v}_j), \qquad k = 1, \dots, K. \tag{5}$$

All fold reconstructions inherit the calibrated world coordinate system of the captured images, so every auxiliary model is queried at the same physical camera. The set $\mathcal{E}_j = \{R_{j,k}\}_{k=1}^{K}$ is the cross-fitted evidence for repair view j. These auxiliary renders need not individually improve upon the full reconstruction. Comparing S_j and F_j with the same references measures how the edit changes their mean disagreement; averaging pools evidence from complementary captured-view subsets.

3.3 Pixel reliability as positive relative risk reduction

For RGB pixels, let $\rho(x, y) = \frac{1}{3}\sum_{c=1}^{3} |x_c - y_c|$ denote channel-averaged absolute distance. For each repair pixel u, we evaluate the initial and repaired candidates against the same fold evidence:

$$B_j(u) = \frac{1}{K}\sum_{k=1}^{K} \rho(S_j(u), R_{j,k}(u)), \tag{6}$$

$$E_j(u) = \frac{1}{K}\sum_{k=1}^{K} \rho(F_j(u), R_{j,k}(u)). \tag{7}$$

Here B_j and E_j measure pre- and post-repair disagreement with the same scene-specific evidence. Their positive relative reduction defines the computable reliability

$$\omega_j(u) = \frac{[B_j(u) - E_j(u)]_+}{B_j(u) + \varepsilon}, \qquad \varepsilon = 10^{-8}, \tag{8}$$

where $[x]_+ = \max(x, 0)$. We average over folds before rectification, so acceptance depends on the signed improvement averaged across folds rather than on a sum of independently rectified per-fold improvements. If the repair does not reduce disagreement, $E_j(u) \geq B_j(u)$ and $\omega_j(u) = 0$; otherwise, the score measures the fraction of pre-repair disagreement removed. Thus $\omega_j(u) \in [0, 1)$ is an empirical reliability score, rather than a calibrated probability of correctness.

3.4 Strictly pixel-wise weighted continuation

After computing ω_j, we resume optimization from the exact sealed main state. Each observed image is assigned an all-one map, whereas each repaired image uses its cross-fitted reliability map. Let C

denote the current render, Y its sampled target, and ω the corresponding weight map. We first retain the spatial form of both components of the standard 3DGS photometric objective:

$$\ell_{\mathrm{L1}}(u) = \frac{1}{3}\sum_{c=1}^{3} |C_c(u) - Y_c(u)|, \tag{9}$$

$$\ell_{\mathrm{DSSIM}}(u) = 1 - \frac{1}{3}\sum_{c=1}^{3} \mathrm{SSIM}_c(C, Y; u). \tag{10}$$

Here $\mathrm{SSIM}_c(C, Y; u)$ is the local SSIM value at pixel u, rather than an image-level scalar. Thus the same $\omega(u)$ acts directly on the L1 and D-SSIM loss maps before either map is reduced.

For a spatial loss map A, we use the fixed-HW (image-normalized) reduction

$$\mathcal{Q}_{\mathrm{img}}(A; \omega) = \frac{1}{HW}\sum_{u\in\Omega} \omega(u)A(u). \tag{11}$$

Writing $\ell_{\mathrm{pixel}}(u) = (1 - \lambda_{\mathrm{ssim}})\ell_{\mathrm{L1}}(u) + \lambda_{\mathrm{ssim}}\ell_{\mathrm{DSSIM}}(u)$, the repair-supervision term abbreviated in Figure 2 is $\mathcal{L}_{\mathrm{rep}} = \mathcal{Q}_{\mathrm{img}}(\ell_{\mathrm{pixel}}; \omega_j)$. Fixed-HW preserves the map magnitude: rejected pixels lower the effective image strength instead of being renormalized by $\sum_u \omega(u)$. For captured images, $\omega \equiv 1$, so Equation 11 is the usual spatial mean.

Using $\lambda_{\mathrm{ssim}} = 0.2$, the continuation objective is

$$\mathcal{L}(C, Y; \omega) = (1 - \lambda_{\mathrm{ssim}})\mathcal{Q}_{\mathrm{img}}(\ell_{\mathrm{L1}}; \omega) + \lambda_{\mathrm{ssim}}\mathcal{Q}_{\mathrm{img}}(\ell_{\mathrm{DSSIM}}; \omega). \tag{12}$$

Equation 12 is a pixel-wise objective: each spatial L1 term and each local D-SSIM term is modulated by its own reliability before aggregation. A single scalar cannot reproduce this gradient unless the entire weight map is constant.

Training protocol. We seal the main 3DGS at the repair step and generate the frozen pairs $\{(S_j, F_j)\}_{j=1}^{N_r}$ once. The K fold models supply evidence for Equation 8; the sealed model then resumes on the combined image pool with Equation 12. Each step samples one full image: real photographs retain unit weights, whereas repaired RGB values and reliability maps remain fixed. Optimizer moments, learning-rate scheduling, and adaptive Gaussian density control continue from the sealed state; fold models are not optimized further.

4 Experiments

4.1 Experimental Setup

Datasets and splits. We use Mip-NeRF 360 (Barron et al., 2022), LLFF (Mildenhall et al., 2019), and Deep Blending (Hedman et al., 2018). Mip-NeRF 360 assigns 3/6/9 images to training, 48 disjoint captured poses to repair, and the remainder to test. LLFF holds out every eighth image, selects 3/6/9 training views, and interpolates 48 repair poses. For Deep Blending, the same holdout and pose-aware farthest traversal give nested 3/6/9 training sets and 48 disjoint repair views. Ground-truth images at repair poses are withheld from training and weighting.

Repair models. We use Difix3D with the nearest-image reference (Wu et al., 2025), and non-temporal, render-only DiffusionHarmonizer (Zhang et al., 2026).

Baselines and metrics. We compare FixTrust-GS with Train-only and uniform whole-image weighting under matched cameras, initialization, optimization, and evaluation. We report PSNR, SSIM (Wang et al., 2004), LPIPS (Zhang et al., 2018), and final Gaussian count (Appendix B.4).

Implementation. Runs share the optimizer and density schedule and use random seed 42 with train-loader seed 10045; seal at 3,000 steps; compute repairs and reliability once; and continue to 9,000 steps with $K = 3$ and fixed-HW. Each uses one 80 GB NVIDIA A100-SXM4 GPU with PyTorch 2.10 and CUDA 12.9. Appendix C.1 specifies the protocol for the detailed Mip-NeRF 360 and Deep Blending studies.

4.2 Main Results

Table 1: Dataset-average results under matched initialization and repair views. Global λ_g is an image weight; FixTrust-GS varies spatially. Train-only enters each repair-model ranking separately; green/blue mark the best/second-best results before rounding.

Dataset	Repair model	Supervision	Train 3 PSNR ↑	Train 3 SSIM ↑	Train 3 LPIPS ↓	Train 6 PSNR ↑	Train 6 SSIM ↑	Train 6 LPIPS ↓	Train 9 PSNR ↑	Train 9 SSIM ↑	Train 9 LPIPS ↓
Mip-NeRF 360	None	Train-only	14.04	0.319	0.545	15.83	0.408	0.441	17.21	0.478	**0.383**
	Difix3D	Global 1.0	14.87	0.340	**0.501**	16.49	0.425	**0.436**	17.62	0.482	0.401
		Global 0.5	14.88	0.346	0.509	16.55	0.432	0.440	17.73	0.491	0.403
		Global 0.3	14.95	0.351	0.513	16.59	0.437	0.444	17.78	0.496	0.405
		FixTrust-GS	**15.18**	**0.361**	0.522	**16.86**	**0.448**	0.445	**18.06**	**0.507**	0.402
	None	Train-only	14.04	0.319	**0.545**	15.83	0.408	**0.441**	17.21	0.478	**0.383**
	Diffusion Harmonizer	Global 1.0	15.34	0.376	0.571	16.93	0.445	0.491	18.02	0.493	0.454
		Global 0.5	15.37	0.380	0.568	16.98	0.450	0.486	18.08	0.500	0.445
		Global 0.3	15.39	0.383	0.566	17.01	0.454	0.480	18.12	0.504	0.437
		FixTrust-GS	**15.49**	**0.387**	0.553	**17.06**	**0.461**	0.466	**18.20**	**0.513**	0.418
Deep Blending	None	Train-only	14.33	0.625	0.663	15.46	0.655	0.609	16.31	0.673	0.573
	Difix3D	Global 1.0	14.77	0.645	0.669	16.03	0.671	0.613	17.02	0.692	0.580
		Global 0.5	14.94	0.653	0.657	16.20	0.677	0.606	17.23	0.700	0.575
		Global 0.3	15.09	0.657	0.653	16.37	0.682	**0.603**	17.26	0.703	**0.573**
		FixTrust-GS	**15.27**	**0.661**	**0.651**	**16.54**	**0.685**	0.603	**17.34**	**0.704**	0.576
	None	Train-only	14.33	0.625	**0.663**	15.46	0.655	**0.609**	16.31	0.673	**0.573**
	Diffusion Harmonizer	Global 1.0	14.77	0.649	0.685	16.05	0.678	0.635	16.88	0.696	0.611
		Global 0.5	14.96	0.655	0.674	16.28	0.683	0.628	**17.11**	0.701	0.603
		Global 0.3	**15.13**	**0.658**	0.669	**16.39**	**0.686**	0.623	17.04	**0.702**	0.600
		FixTrust-GS	15.06	0.657	0.665	16.19	0.684	0.622	16.99	0.702	0.593
LLFF	None	Train-only	18.37	0.593	0.253	**21.45**	**0.704**	0.189	**22.02**	**0.729**	**0.180**
	Difix3D	Global 1.0	19.53	0.630	**0.219**	21.30	0.699	**0.188**	21.53	0.709	0.189
		Global 0.5	19.56	0.630	0.223	21.32	0.700	0.191	21.59	0.712	0.192
		Global 0.3	19.42	0.623	0.231	21.26	0.696	0.196	21.51	0.710	0.196
		FixTrust-GS	**19.66**	**0.633**	0.227	21.39	0.704	0.195	21.66	0.718	0.196
	None	Train-only	18.37	0.593	0.253	21.45	0.704	0.189	**22.02**	**0.729**	**0.180**
	Diffusion Harmonizer	Global 1.0	19.48	0.625	0.224	21.25	0.692	0.193	21.52	0.704	0.194
		Global 0.5	19.61	0.632	0.219	21.38	0.705	0.188	21.63	0.715	0.189
		Global 0.3	19.39	0.620	0.228	21.18	0.688	0.199	21.45	0.700	0.200
		FixTrust-GS	**19.82**	**0.641**	**0.216**	**21.52**	**0.712**	**0.185**	21.80	0.728	0.186

Aggregate reconstruction quality. Table 1 compares matched supervision. FixTrust-GS gives the strongest repaired-supervision PSNR and SSIM in most blocks, supporting spatial validation across models. Global weighting remains competitive for DiffusionHarmonizer on Deep Blending, and Train-only on LLFF; LPIPS varies more because repair can alter captured appearance. Cases in Appendix C.4 cover both repair models on Mip-NeRF 360 and Deep Blending at each view budget.

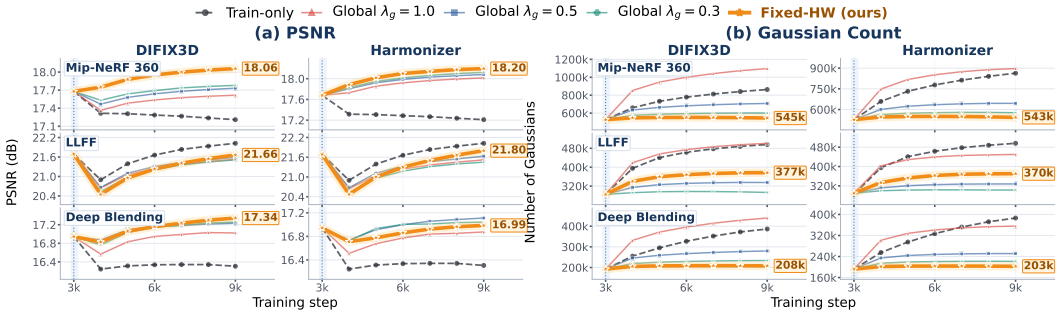

Figure 3: Dataset-average trajectories at Train 9 ($N_r = 48$). Repairs enter after the dotted 3k-step boundary; orange labels give final FixTrust-GS PSNR and Gaussian counts (thousands).

Optimization dynamics. Figure 3(a) shows an initial perturbation followed by recovery. FixTrust-GS leads the repaired PSNR trajectories at the endpoint on Mip-NeRF 360 and LLFF; on Deep Blending, the ranking depends on the repair model.

Table 2: Scene-wise Mip-NeRF 360 results at Train 9 with Difix3D (#GS in thousands).

Method	Metric	Bicycle	Bonsai	Counter	Flowers	Garden	Kitchen	Room	Stump	Treehill
TRAIN-ONLY	PSNR ↑	15.17	17.62	17.92	12.83	20.93	20.80	19.17	16.75	13.69
	SSIM ↑	0.305	0.598	0.555	0.248	0.611	0.719	0.649	0.273	0.342
	LPIPS ↓	**0.527**	0.311	0.290	**0.554**	**0.266**	0.161	0.293	**0.474**	**0.571**
	#GS (10^3) ↓	1487 (ref.)	629 (ref.)	300 (ref.)	1150 (ref.)	802 (ref.)	473 (ref.)	370 (ref.)	1446 (ref.)	1111 (ref.)
Global 1.0	PSNR ↑	15.82	17.95	18.10	13.41	21.21	20.94	19.09	17.45	14.57
	SSIM ↑	0.321	0.609	0.561	0.257	0.594	0.703	0.646	0.285	0.366
	LPIPS ↓	0.550	0.305	0.296	0.586	0.310	0.194	0.292	0.499	0.573
	#GS (10^3) ↓	**1633** (↑9.9%)	**1010** (↑60.5%)	**532** (↑77.3%)	**1333** (↑15.9%)	**1048** (↑30.6%)	**703** (↑48.6%)	**567** (↑53.2%)	**1849** (↑27.9%)	**1199** (↑7.9%)
Global 0.5	PSNR ↑	15.88	18.14	18.20	13.48	21.29	21.18	19.23	17.54	14.65
	SSIM ↑	0.329	0.621	0.568	0.263	0.601	0.719	0.653	0.294	0.372
	LPIPS ↓	0.561	0.297	0.294	0.595	0.315	0.180	0.290	0.511	0.583
	#GS (10^3) ↓	1080 (↓27.3%)	679 (↑7.8%)	342 (↑14.0%)	844 (↓26.6%)	647 (↓19.3%)	497 (↑5.1%)	378 (↑2.0%)	1158 (↓19.9%)	742 (↓33.2%)
Global 0.3	PSNR ↑	15.92	18.21	18.26	13.48	21.36	21.26	19.31	17.58	14.62
	SSIM ↑	0.333	0.626	0.573	0.266	0.605	0.727	0.657	0.299	0.375
	LPIPS ↓	0.569	0.295	0.291	0.602	0.316	0.172	0.288	0.520	0.592
	#GS (10^3) ↓	925 (↓37.8%)	586 (↓6.8%)	290 (↓3.4%)	705 (↓38.7%)	553 (↓31.0%)	445 (↓5.9%)	324 (↓12.6%)	955 (↓33.9%)	619 (↓44.3%)
FIXTRUST-GS	PSNR ↑	**16.15**	**18.72**	**18.64**	**13.70**	**21.53**	**21.43**	**19.82**	**17.68**	**14.85**
	SSIM ↑	**0.343**	**0.639**	**0.586**	**0.274**	**0.614**	**0.741**	**0.673**	**0.309**	**0.382**
	LPIPS ↓	0.571	**0.281**	**0.283**	0.608	0.312	**0.152**	**0.279**	0.534	0.598
	#GS (10^3) ↓	**847** (↓43.0%)	**542** (↓13.8%)	**264** (↓12.2%)	**626** (↓45.5%)	**511** (↓36.3%)	**418** (↓11.6%)	**291** (↓21.4%)	**849** (↓41.3%)	**556** (↓50.0%)

Gaussian efficiency. The gains of FIXTRUST-GS do not require a larger Gaussian representation. At Train 9, it improves PSNR and SSIM across Mip-NeRF 360 with 544.9 thousand Gaussians, 36.9% fewer than Train-only and fewer in every scene (11.6–50.0%). Global 1.0 instead enlarges every representation, reaching 1,097.1 thousand Gaussians—27.1% above Train-only and over twice the size of FIXTRUST-GS. Together with Figure 3(b), this is consistent with uniform repair gradients preserving unsupported residuals and prompting densification, while reliability weighting limits those gradients without discarding supported repair evidence.

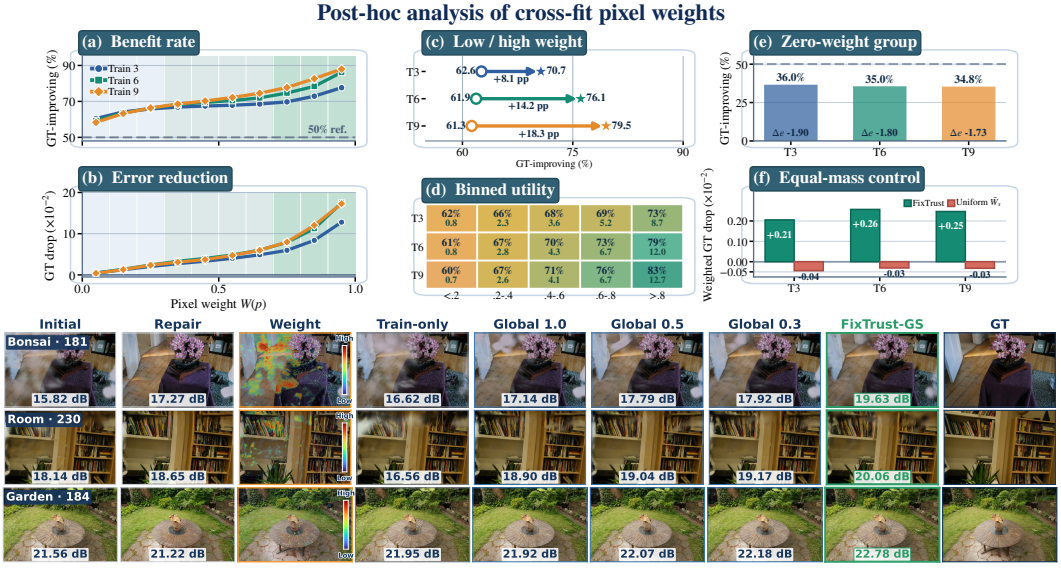

Figure 4: **Post-hoc weight reliability and qualitative cases.** Top: equal-scene Difix3D ground-truth diagnostics at $N_r = 48$ across Mip-NeRF 360; Appendix C.2.1 gives scope and definitions. Panels (a–e) stratify ground-truth repair benefit; (f) compares the cross-fitted reliability map with the equal-mass uniform scene mean $\bar{W}_s$ over all repaired-view pixels. Bottom: same-camera inputs, overlays, and final renders with PSNR.

Post-hoc weight reliability. Ground truth serves as a post-hoc reference; Appendix C.2.1 defines the metrics, scope, and equal-scene aggregation. Across Train 3/6/9, benefit frequency rises from 58.4–60.5% in the lowest active bin to 77.6–88.0% in the highest, while only 34.8–36.0% of zero-weight pixels improve and their mean reduction is negative. With scene-level mass fixed, cross-fitted reliability maps yield $+0.21$–$+0.26$ weighted ground-truth error reduction versus -0.04–-0.03 for uniform $\bar{W}_s$. This offline gain reflects weight allocation both between and within repaired views.

4.3 Ablation Studies

Global repair strength. Table 1 shows that no global weight is consistently best across datasets, view budgets, repair models, and metrics. Dense weights adapt supervision to local evidence. Appendix C.3 provides the full scene-wise sweeps over repair-view counts and scalar weights.

The active-pixel reduction used in the ablation is

$$\mathcal{Q}_{\mathrm{act}}(A;\omega) = \frac{\sum_{u\in\Omega}\omega(u)A(u)}{\max\left(\sum_{u\in\Omega}\mathbf{1}[\omega(u)>0],1\right)}. \tag{13}$$

Loss-reduction control. Equation 13 averages the accepted pixels: the numerator weights each residual by its reliability, while the denominator counts only pixels with positive weight. It therefore removes the area penalty of sparse support. Fixed-HW instead keeps the full HW denominator, so the repaired view contributes in proportion to both confidence and reliable area.

Cross-fitting. Table 3 and Figure 5(a) show that $K = 3$ gives the best PSNR for both repair models. Larger K changes the other metrics only marginally while increasing auxiliary cost, so we use three pose-balanced folds throughout.

Ablation results. Figure 5(b) fixes the evidence and varies supervision control. Spatial weights outperform global weighting. Table 3 shows that active-pixel favors PSNR; fixed-HW improves SSIM and reduces Gaussians for both models, with model-dependent LPIPS changes. Similar $K = 5/9$ results suggest saturation. Appendix C.2 compares the reliability formulas.

Table 3: K-fold and loss-reduction ablations on Mip-NeRF 360 (Train 9, $N_r = 48$). Fold rows use fixed-HW; reduction rows use $K = 3$. Bold: defaults; colors: top two.

Repair model	Setting	PSNR ↑	SSIM ↑	LPIPS ↓	Gaussians (10^3) ↓ change vs. Train-only
Fold count (fixed-HW)					
Difix3D	$K = 2$	18.00	0.504	0.404	547.8 (↓36.5%)
	K = 3	**18.06**	0.507	**0.402**	544.9 (↓36.9%)
	$K = 5$	18.02	0.507	0.402	543.8 (↓37.0%)
	$K = 9$	18.03	**0.508**	0.402	**542.7** (↓37.1%)
Diffusion Harmonizer	$K = 2$	18.14	0.510	0.421	543.8 (↓36.9%)
	K = 3	**18.20**	0.513	0.418	542.9 (↓37.0%)
	$K = 5$	18.19	0.514	0.416	541.2 (↓37.2%)
	$K = 9$	18.16	0.514	0.416	**541.1** (↓37.3%)
Loss reduction ($K = 3$)					
Difix3D	Active-pixel	**18.08**	0.504	**0.400**	609.5 (↓29.4%)
	Fixed-HW	18.06	**0.507**	0.402	**544.9** (↓36.9%)
Diffusion Harmonizer	Active-pixel	**18.24**	0.511	0.426	573.6 (↓33.5%)
	Fixed-HW	18.20	**0.513**	**0.418**	**542.9** (↓37.0%)

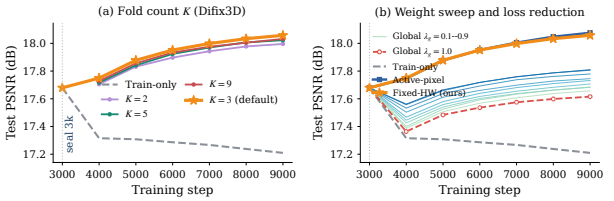

Figure 5: Mip-NeRF 360 ablations; repair at 3k steps.

5 Conclusion

We introduced FixTrust-GS to incorporate repaired novel views into sparse-view 3DGS through pixel-wise evidence from the captured scene. Cross-fitted reconstructions compare the original and repaired pixels at the same camera, producing dense supervision weights for outputs from different repair models. Experiments on multiple datasets with two repair models and three sparse-view budgets show that FixTrust-GS improves reconstruction over uniform image-level weighting, with broad gains in PSNR and SSIM and competitive LPIPS. The resulting models also use fewer Gaussian primitives, suggesting that rejecting unsupported repair content reduces the representation capacity spent fitting hallucinated structure.

Limitations and Future Work. Cross-fitted predictions may share errors in regions that remain weakly constrained by the captured views, while pixel-wise evidence does not fully capture three-dimensional geometry or semantic consistency. Future work will combine the proposed evidence with depth, correspondence, and semantic cues, and study calibrated uncertainty across repair models, view distributions, and scene types.

AI use statement. Generative AI supported (i) writing and language refinement, (ii) literature discovery, and (iii) research planning and implementation. The authors verified assisted content against experiments and source materials and take responsibility for the manuscript. AI neither generated reported measurements nor supplied synthetic results for missing experiments.

References

Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, and Peter Hedman. Mip-NeRF 360: Unbounded anti-aliased neural radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5470–5479, 2022.

David Charatan, Sizhe Lester Li, Andrea Tagliasacchi, and Vincent Sitzmann. pixelsplat: 3d gaussian splats from image pairs for scalable generalizable 3d reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 19457–19467, 2024.

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 (ECCV)*, 2024.

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 *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 27126–27136, 2025.

Alex Hanson, Allen Tu, Vasu Singla, Mayuka Jayawardhana, Matthias Zwicker, and Tom Goldstein. PUP 3D-GS: Principled uncertainty pruning for 3d gaussian splatting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5949–5958, 2025.

Peter Hedman, Julien Philip, True Price, Jan-Michael Frahm, George Drettakis, and Gabriel J. Brostow. Deep blending for free-viewpoint image-based rendering. *ACM Transactions on Graphics*, 37(6):257:1–257:15, 2018. doi: 10.1145/3272127.3275084.

Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2d gaussian splatting for geometrically accurate radiance fields. In *ACM SIGGRAPH Conference Papers*, 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 *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 688–697, 2026.

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4):139:1–139:14, 2023. doi: 10.1145/3592433.

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

Jonas Kulhanek, Songyou Peng, Zuzana Kukelova, Marc Pollefeys, and Torsten Sattler. WildGaussians: 3d gaussian splatting in the wild. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 37, 2024.

Jiahe Li, Jiawei Zhang, Xiao Bai, Jin Zheng, Xin Ning, Jun Zhou, and Lin Gu. DNGaussian: Optimizing sparse-view 3d gaussian radiance fields with global-local depth normalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20775–20785, 2024.

Xi Liu, Chaoyi Zhou, and Siyu Huang. 3DGS-Enhancer: Enhancing unbounded 3d gaussian splatting with view-consistent 2d diffusion priors. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 37, 2024. doi: 10.52202/079017-4237.

Xi Liu, Weiwei Sun, Zhou Ren, Chris Broaddus, Siyu Huang, and Laurent Guigues. HAD: Hallucination-aware diffusion priors for 3d reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 29781–29791, 2026.

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*, 38(4):29:1–29:14, 2019. doi: 10.1145/3306346.3322980.

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 on Computer Vision (ECCV)*, 2020.

Michael Niemeyer, Jonathan T. Barron, Ben Mildenhall, Mehdi S. M. Sajjadi, Andreas Geiger, and Noha Radwan. RegNeRF: Regularizing neural radiance fields for view synthesis from sparse inputs. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5480–5490, 2022.

Sara Sabour, Lily Goli, George Kopanas, Mark Matthews, Dmitry Lagun, Leonidas Guibas, Alec Jacobson, David J. Fleet, and Andrea Tagliasacchi. SpotLessSplats: Ignoring distractors in 3d gaussian splatting. *arXiv preprint arXiv:2406.20055*, 2024.

Luca Savant Aira, Diego Valsesia, and Enrico Magli. Modeling uncertainty for gaussian splatting. *IEEE Transactions on Neural Networks and Learning Systems*, 36(6):11657–11663, 2025.

Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In *International Conference on Learning Representations (ICLR)*, 2015.

Rui Song, Tianhui Cai, Markus Gross, Xingcheng Zhou, Zewei Zhou, Zhiyu Huang, Olaf Wysocki, and Jiaqi Ma. ConFixGS: Learning to fix feedforward 3d gaussian splatting with confidence-aware diffusion priors in driving scenes. *arXiv preprint arXiv:2605.09688*, 2026. URL `https://arxiv.org/abs/2605.09688`.

Atakan Topaloğlu, Kunyi Li, Michael Niemeyer, Nassir Navab, A. Murat Tekalp, and Federico Tombari. OracleGS: Grounding generative priors for sparse-view gaussian splatting. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 77–87, 2026.

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

YuAn Wang, Xiaofan Li, Chi Huang, Wenhao Zhang, Hao Li, Bosheng Wang, Xun Sun, and Jun Wang. FaithFusion: Harmonizing reconstruction and generation via pixel-wise information gain. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1198–1209, 2026.

Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600–612, 2004. doi: 10.1109/TIP.2003.819861.

Jiaxin Wei, Stefan Leutenegger, and Simon Schaefer. GSFix3D: Diffusion-guided repair of novel views in gaussian splatting. In *International Conference on 3D Vision (3DV)*, pp. 1051–1060, 2026. doi: 10.1109/3DV69130.2026.00105.

Jay Zhangjie Wu, Yuxuan Zhang, Haithem Turki, Xuanchi Ren, Jun Gao, Mike Zheng Shou, Sanja Fidler, Zan Gojcic, and Huan Ling. DIFIX3D+: Improving 3d reconstructions with single-step diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 26024–26035, 2025.

Rundi Wu, Ben Mildenhall, Philipp Henzler, Keunhong Park, Ruiqi Gao, Daniel Watson, Pratul P. Srinivasan, Dor Verbin, Jonathan T. Barron, Ben Poole, and Aleksander Holynski. ReconFusion: 3d reconstruction with diffusion priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 21551–21561, 2024.

Jamie Wynn and Daniyar Turmukhambetov. DiffusioNeRF: Regularizing neural radiance fields with denoising diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4180–4189, 2023.

Haolin Xiong, Sairisheek Muttukuru, Hanyuan Xiao, Rishi Upadhyay, Pradyumna Chari, Yajie Zhao, and Achuta Kadambi. SparseGS: Sparse view synthesis using 3d gaussian splatting. In *International Conference on 3D Vision (3DV)*, pp. 1032–1041, 2025. doi: 10.1109/3DV66043.2025.00100.

Haofei Xu, Songyou Peng, Fangjinhua Wang, Hermann Blum, Daniel Barath, Andreas Geiger, and Marc Pollefeys. DepthSplat: Connecting gaussian splatting and depth. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 16453–16463, 2025a.

Yexing Xu, Longguang Wang, Minglin Chen, Sheng Ao, Li Li, and Yulan Guo. DropoutGS: Dropping out gaussians for better sparse-view rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 701–710, 2025b.

Jiawei Yang, Marco Pavone, and Yue Wang. FreeNeRF: Improving few-shot neural rendering with free frequency regularization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8254–8263, 2023.

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

Jiawei Zhang, Jiahe Li, Xiaohan Yu, Lei Huang, Lin Gu, Jin Zheng, and Xiao Bai. CoR-GS: Sparse-view 3d gaussian splatting via co-regularization. In *European Conference on Computer Vision (ECCV)*, pp. 335–352, 2024.

Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 586–595, 2018.

Yuxuan Zhang, Katarína Tóthová, Zian Wang, Kangxue Yin, Haithem Turki, Riccardo de Lutio, Yen-Yu Chang, Or Litany, Sanja Fidler, and Zan Gojcic. DiffusionHarmonizer: Bridging neural reconstruction and photorealistic simulation with online diffusion enhancer. *arXiv preprint arXiv:2602.24096*, 2026. URL https://arxiv.org/abs/2602.24096.

Yingji Zhong, Zhihao Li, Dave Zhenyu Chen, Lanqing Hong, and Dan Xu. Taming video diffusion prior with scene-grounding guidance for 3d gaussian splatting from sparse inputs. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6133–6143, 2025.

Hongyu Zhou, Zisen Shao, Sheng Miao, Pan Wang, Dongfeng Bai, Bingbing Liu, and Yiyi Liao. FreeFix: Boosting 3d gaussian splatting via fine-tuning-free diffusion models. In *International Conference on 3D Vision (3DV)*, pp. 1091–1106, 2026. doi: 10.1109/3DV69130.2026.00109. URL https://arxiv.org/abs/2601.20857.

Zehao Zhu, Zhiwen Fan, Yifan Jiang, and Zhangyang Wang. FSGS: Real-time few-shot view synthesis using gaussian splatting. In *European Conference on Computer Vision (ECCV)*, 2024.

Appendix

Where Fixes Really Help:

Cross-Fitted Repair-Model Supervision for 3DGS

Contents

Appendix A Notation **14**

A.1 Pose-balanced fold construction 15

Appendix B 3D Gaussian Splatting Preliminaries **16**

B.1 Scene representation 16

B.2 Projection and compositing 16

B.3 Photometric optimization 16

B.4 Evaluation metrics 16

Appendix C Extended Results Across Repair Models and Datasets **17**

C.1 Evaluation matrix and shared protocol 17

C.2 Shared ablation controls 17

C.2.1 Post-hoc diagnostic metrics 18

C.3 Full quantitative results 18

C.3.1 Difix3D on Mip-NeRF 360 18

C.3.2 Difix3D on Deep Blending 23

C.3.3 Diffusion Harmonizer on Mip-NeRF 360 27

C.3.4 Diffusion Harmonizer on Deep Blending 31

C.4 Representative qualitative comparisons 35

Difix3D on Mip-NeRF 360, Train 3/6/9 35–37

Difix3D on Deep Blending, Train 3/6/9 38–40

Diffusion Harmonizer on Mip-NeRF 360, Train 3/6/9 41–43

Diffusion Harmonizer on Deep Blending, Train 3/6/9 44–46

A Notation

Table 4 consolidates the notation used by the method, the pose-balanced split, and the 3DGS review. Images and renderings are normalized to $[0,1]^{H\times W\times 3}$ unless stated otherwise.

Table 4: Notation used throughout the paper. A superscript $-k$ denotes exclusion of fold k; $[x]_+ = \max(x, 0)$ and $\mathbf{1}[\cdot]$ is an indicator.

Symbol	Domain or size	Meaning
Indices, counts, and image domain		
n, j, k, i	Integer indices	Captured view, repair view, cross-fitting fold, and Gaussian primitive.
u, c	$u \in \Omega$, $c \in \{1,2,3\}$	Pixel location and RGB channel.
N_t, N_r, K, M	Positive integers	Numbers of captured views, repaired views, folds, and Gaussian primitives.
H, W, Ω	$\Omega = \{1{:}H\} \times \{1{:}W\}$	Image height, width, and pixel domain; $\lvert\Omega\rvert = HW$.
Captured and repaired supervision		
v_n, I_n	Camera; $[0,1]^{H\times W\times 3}$	Captured camera and its observed RGB image.
$\mathcal{T}$	$\{(v_n, I_n)\}_{n=1}^{N_t}$	Calibrated captured-view training set.
$\widetilde{v}_j, \mathcal{V}_r$	Camera; $\{\widetilde{v}_j\}_{j=1}^{N_r}$	Repair camera and the complete repair-camera set.
$\theta, \mathcal{R}(\theta; v)$	Parameters; RGB image	3DGS parameters and their differentiable rendering from camera v.
$\theta^0, \theta^\star$	3DGS parameters	Sealed reconstruction before repair and final state after continuation.
S_j	RGB image	Initial render $\mathcal{R}(\theta^0; \widetilde{v}_j)$ supplied to the repair model.
$\mathrm{Fix}_\phi(\cdot; \xi_j)$	Pretrained mapping	Repair model with fixed parameters ϕ and optional condition ξ_j.
F_j	RGB image	Repaired output $\mathrm{Fix}_\phi(S_j; \xi_j)$ at camera $\widetilde{v}_j$.
Cross-fitted evidence and pose-balanced folds		
$\mathcal{H}_k$	Index subset	Captured-view indices held out from auxiliary model k.
$\mathcal{T}_{-k}, \theta^{-k}$	Training set; parameters	Complement of $\mathcal{H}_k$ and the auxiliary 3DGS trained on it.
$R_{j,k}$	RGB image	Auxiliary render $\mathcal{R}(\theta^{-k}; \widetilde{v}_j)$ used as evidence.
$\mathcal{E}_j$	$\{R_{j,k}\}_{k=1}^{K}$	Cross-fitted evidence ensemble for repair view j.
$\mathbf{o}_n, \mathbf{d}_n$	$\mathbb{R}^3$	Camera center and unit forward direction of captured view n.
$\bar{\mathbf{o}}, s_o, \mathbf{z}_n$	$\mathbb{R}^3$; $\mathbb{R}_+$; $\mathbb{R}^6$	Mean camera center, positional scale, and normalized pose descriptor.
D_{nm}	$\mathbb{R}_{\geq 0}$	Euclidean distance $\lVert\mathbf{z}_n - \mathbf{z}_m\rVert_2$ between pose descriptors.
s_k, r_k	Integer; $\mathbb{R}_{\geq 0}$	Fold size and coverage radius of fold k.
$\bar{r}, \sigma_r, q_k, q$	$\mathbb{R}_{\geq 0}$	Mean/spread of fold coverage and within-fold separation statistics.
Pixel reliability and weighted continuation		
$\rho(x, y)$	$[0,1]^3 \times [0,1]^3 \to \mathbb{R}_{\geq 0}$	Channel-averaged absolute distance between RGB pixels.
$B_j(u)$	$\mathbb{R}_{\geq 0}$	Pre-repair disagreement between $S_j(u)$ and the evidence ensemble.
$E_j(u)$	$\mathbb{R}_{\geq 0}$	Post-repair disagreement between $F_j(u)$ and the same evidence ensemble.
$\omega_j(u)$	$[0, 1)$	Pixel reliability given by the positive relative reduction from $B_j(u)$ to $E_j(u)$.
ε	10^{-8}	Numerical stabilizer in the reliability denominator.
C, Y, ω	RGB, RGB, scalar map	Generic render, supervision target, and spatial weight during continuation.
$\ell_{\mathrm{L1}}, \ell_{\mathrm{DSSIM}}$	$\Omega \to \mathbb{R}_{\geq 0}$	Per-pixel photometric and structural-dissimilarity maps.
ℓ_{pixel}	$\Omega \to \mathbb{R}_{\geq 0}$	Weighted per-pixel combination of the two loss maps.

Continued on the next page.

Table 4 continued.

Symbol	Domain or size	Meaning
$\mathcal{Q}_{\mathrm{img}}(A;\omega)$	Scalar	Fixed-image-area reduction of map A under spatial weights ω.
$\mathcal{L}_{\mathrm{rep}}$	Scalar	Fixed-HW repair loss $\mathcal{Q}_{\mathrm{img}}(\ell_{\mathrm{pixel}};\omega_j)$.
$\lambda_{\mathrm{ssim}}, \lambda_g$	$[0,1]$	Structural-loss mixing coefficient and scalar whole-image repair weight.
3D Gaussian Splatting		
$G_i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i$	Kernel; $\mathbb{R}^3$; $\mathbb{S}^3_{++}$	Gaussian primitive, center, and 3D covariance.
$\alpha_i, \mathbf{h}_i, \mathbf{c}_i(v)$	$[0,1]$; SH coefficients; RGB	Opacity, spherical-harmonic coefficients, and view-dependent color.
$\mathbf{R}_i, \mathbf{S}_i, \mathbf{s}_i$	$\mathrm{SO}(3)$; diagonal; $\mathbb{R}^3_{>0}$	Rotation, scale matrix, and principal scales used to parameterize $\boldsymbol{\Sigma}_i$.
$\pi_v, \mathbf{W}_v, \mathbf{J}_{i,v}$	Projection and matrices	Camera projection, world-to-camera linear map, and local projection Jacobian.
$\boldsymbol{\mu}^{2D}_{i,v}, \boldsymbol{\Sigma}^{2D}_{i,v}$	$\mathbb{R}^2$; $\mathbb{S}^2_{++}$	Projected mean and covariance of primitive i in camera v.
$\alpha_{i,v}(u), T_{i,v}(u)$	$[0,1]$	Effective pixel opacity and accumulated transmittance in depth order.
$\mathbf{C}_v(u)$	$[0,1]^3$	Alpha-composited 3DGS color at pixel u from camera v.
$\mathcal{L}_{\mathrm{3DGS}}$	Scalar	Standard captured-view 3DGS reconstruction objective.
Evaluation metrics		
$\widehat{I}, I^{\mathrm{gt}}$	RGB images	Predicted image and held-out ground truth.
MSE, PSNR	Scalars	Mean-squared error and peak signal-to-noise ratio.
SSIM, LPIPS	Scalars	Structural-similarity score and learned perceptual distance.
Δ^{gt}_q	Scalar	Signed reduction in per-pixel ground-truth error after repair; positive means improvement.
$P_s(\mathcal{A}), D_s(\mathcal{A})$	Scalars	Benefit rate and mean ground-truth error reduction on pixel set $\mathcal{A}$ in scene s.
$\overline{W}_s, G^{\mathrm{map}}_s, G^{\mathrm{uniform}}_s$	Scalars	Mean weight and equal-mass weighted ground-truth error reductions for the cross-fitted reliability and uniform maps.

A.1 Pose-balanced fold construction

To make the held-out evidence cover the training trajectory, we form the pose descriptor and pairwise distance

$$\mathbf{z}_n = \left[\frac{\mathbf{o}_n - \bar{\mathbf{o}}}{s_o};\, \mathbf{d}_n\right], \qquad s_o = \left(\frac{1}{N_t}\sum_{n=1}^{N_t}\|\mathbf{o}_n - \bar{\mathbf{o}}\|_2^2\right)^{1/2}, \qquad D_{nm} = \|\mathbf{z}_n - \mathbf{z}_m\|_2, \tag{14}$$

where $\bar{\mathbf{o}}$ is the mean camera center. We restrict each held-out fold to $s_k \in \{\lfloor N_t/K \rfloor, \lceil N_t/K \rceil\}$ and score a partition by its coverage radius

$$r_k = \max_{n\in\{1,\dots,N_t\}} \min_{m\in\mathcal{H}_k} D_{nm}, \qquad \bar{r} = \frac{1}{K}\sum_{k=1}^{K} r_k, \qquad \sigma_r = \left(\frac{1}{K}\sum_{k=1}^{K}(r_k - \bar{r})^2\right)^{1/2}. \tag{15}$$

The selected partition lexicographically minimizes

$$\left(\max_k r_k,\ \bar{r},\ \sigma_r,\ -q\right), \qquad q = \min_k q_k, \qquad q_k = \begin{cases} \min_{\substack{m,m'\in\mathcal{H}_k \\ m\neq m'}} D_{mm'}, & |\mathcal{H}_k| \geq 2, \\ +\infty, & |\mathcal{H}_k| = 1, \end{cases} \tag{16}$$

which first equalizes worst-case coverage, then average coverage and its spread, and finally favors separation within each fold. We enumerate feasible partitions for small splits and use deterministic greedy assignment followed by pairwise swaps for larger splits.

B 3D Gaussian Splatting Preliminaries

B.1 Scene representation

A 3DGS scene contains M anisotropic Gaussian primitives $\{G_i\}_{i=1}^{M}$. Following the original formulation (Kerbl et al., 2023), the spatial influence of primitive i is

$$G_i(\mathbf{x}) = \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathsf{T}}\boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)\right], \tag{17}$$

where $\boldsymbol{\mu}_i \in \mathbb{R}^3$ and $\boldsymbol{\Sigma}_i \in \mathbb{R}^{3\times 3}$ are its mean and covariance. Each primitive additionally stores an opacity $o_i \in [0,1]$ and spherical-harmonic coefficients $\mathbf{h}_i$, which produce the view-dependent RGB color $\mathbf{c}_i(v)$. To maintain a valid covariance, 3DGS parameterizes

$$\boldsymbol{\Sigma}_i = \mathbf{R}_i\mathbf{S}_i\mathbf{S}_i^{\mathsf{T}}\mathbf{R}_i^{\mathsf{T}}, \qquad \mathbf{S}_i = \mathrm{diag}(\mathbf{s}_i), \tag{18}$$

where $\mathbf{s}_i \in \mathbb{R}^3_{>0}$ stores the principal scales and $\mathbf{R}_i \in \mathrm{SO}(3)$ is the rotation matrix of a unit quaternion.

B.2 Projection and compositing

For camera v, let π_v denote perspective projection, $\mathbf{W}_v$ the linear part of its world-to-camera transform, and $\mathbf{J}_{i,v}$ the Jacobian of the projective map at $\boldsymbol{\mu}_i$. The projected mean and covariance are

$$\boldsymbol{\mu}'_{i,v} = \pi_v(\boldsymbol{\mu}_i), \qquad \boldsymbol{\Sigma}'_{i,v} = \mathbf{J}_{i,v}\mathbf{W}_v\boldsymbol{\Sigma}_i\mathbf{W}_v^{\mathsf{T}}\mathbf{J}_{i,v}^{\mathsf{T}}. \tag{19}$$

Evaluating the projected Gaussian at pixel $u \in \Omega$ gives its effective alpha value

$$\alpha_{i,v}(u) = o_i \exp\left[-\frac{1}{2}(u-\boldsymbol{\mu}'_{i,v})^{\mathsf{T}}(\boldsymbol{\Sigma}'_{i,v})^{-1}(u-\boldsymbol{\mu}'_{i,v})\right]. \tag{20}$$

After sorting the visible primitives by depth, the rasterizer uses front-to-back alpha compositing,

$$\mathbf{C}_v(u) = \sum_{i=1}^{M} T_{i,v}(u)\alpha_{i,v}(u)\mathbf{c}_i(v), \qquad T_{i,v}(u) = \prod_{q<i}\big(1-\alpha_{q,v}(u)\big). \tag{21}$$

Evaluating only primitives that overlap the pixel tile enables real-time differentiable rendering.

B.3 Photometric optimization

Given captured views $\mathcal{T} = \{(v_n, I_n)\}_{n=1}^{N_t}$, standard 3DGS jointly optimizes primitive positions, covariances, opacities, and colors using

$$\mathcal{L}_{\mathrm{3DGS}} = (1-\lambda_{\mathrm{ssim}})\mathcal{L}_1(\mathbf{C}_{v_n}, I_n) + \lambda_{\mathrm{ssim}}\big(1-\mathrm{SSIM}(\mathbf{C}_{v_n}, I_n)\big), \tag{22}$$

where $\mathcal{L}_1$ is the mean absolute RGB error. Optimization is interleaved with adaptive densification and pruning. Under sparse views, this objective constrains only the observed rays; primitives in weakly observed regions can therefore explain the training images while producing incorrect novel views. In FixTrust-GS, repaired-image losses use the cross-fitted weights in Equation 12, while captured images retain unit weight.

B.4 Evaluation metrics

Let $\widehat{I}, I^{\mathrm{gt}} \in [0,1]^{H\times W\times 3}$ denote a rendered test image and its held-out ground truth. We use the same three full-reference metrics as the original 3DGS evaluation protocol. Mean squared error and peak signal-to-noise ratio are

$$\mathrm{MSE}(\widehat{I}, I^{\mathrm{gt}}) = \frac{1}{3HW}\|\widehat{I}-I^{\mathrm{gt}}\|_F^2, \qquad \mathrm{PSNR}(\widehat{I}, I^{\mathrm{gt}}) = 10\log_{10}\frac{L_{\max}^2}{\mathrm{MSE}(\widehat{I}, I^{\mathrm{gt}})}, \tag{23}$$

where $L_{\max} = 1$ for normalized RGB images. Higher PSNR indicates lower pixel-wise squared error.

For a corresponding pair of local image windows x and y, structural similarity is (Wang et al., 2004)

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \tag{24}$$

where μ_x, μ_y are local means, σ_x^2, σ_y^2 are local variances, and σ_{xy} is the local covariance. Following the 3DGS implementation, we use an 11×11 Gaussian window with standard deviation 1.5, $C_1 = (0.01L_{\max})^2$, and $C_2 = (0.03L_{\max})^2$, then average over pixels and RGB channels. Higher SSIM indicates greater structural agreement.

LPIPS compares normalized deep features rather than RGB values (Zhang et al., 2018). Let $\widehat{\mathbf{z}}_{l,h,w}(I)$ denote the unit-normalized activation at layer l and location (h, w), and let $\mathbf{w}_l$ be its learned channel calibration. The image distance is

$$\mathrm{LPIPS}(\widehat{I}, I^{\mathrm{gt}}) = \sum_{l \in \mathcal{L}} \frac{1}{H_l W_l} \sum_{h=1}^{H_l} \sum_{w=1}^{W_l} \left\| \mathbf{w}_l \odot \left(\widehat{\mathbf{z}}_{l,h,w}(\widehat{I}) - \widehat{\mathbf{z}}_{l,h,w}(I^{\mathrm{gt}}) \right) \right\|_2^2. \tag{25}$$

We use the calibrated VGG variant (Simonyan & Zisserman, 2015), matching the public 3DGS evaluator; lower LPIPS indicates closer perceptual similarity. Each metric is first averaged over the held-out views of a scene, and dataset-level entries report the unweighted mean across scenes.

C Extended Results Across Repair Models and Datasets

C.1 Evaluation Matrix and Shared Protocol

We organize the detailed results as the repair-model–dataset matrix in Table 5. Every completed cell follows the same reporting order: aggregate response, exact Train 3/6/9 values, scene-level results, and a concise interpretation. Representative aligned qualitative comparisons follow the quantitative matrix. This common structure keeps the appendix focused on results and makes the cells directly comparable.

Table 5: Repair-model–dataset matrix used by the extended results. Every cell evaluates Train 3/6/9 and 48 repaired views in the main comparison; detailed modules additionally report all available repair-view and weight sweeps.

Repair model	Mip-NeRF 360	Deep Blending
Difix3D	Train 3/6/9	Train 3/6/9
Diffusion Harmonizer	Train 3/6/9	Train 3/6/9

Shared protocol. Both datasets report nested $N_t \in \{3, 6, 9\}$ subsets, $N_r = 48$, and global weights $\lambda_g \in \{0.3, 0.5, 1.0\}$ in the main comparison. All methods use matched repair cameras and evaluation views and resume from the same step-3,000 checkpoint to step 8,999. Repair outputs and cross-fitted evidence are cached once. Runs use seed 42, data-loader seed 10045, $K = 3$ pose-balanced folds, and fixed-HW reduction; dataset values are equal-scene macro-averages. Captured images retain unit supervision, so comparisons change only how the same repaired RGB enters continuation.

C.2 Shared Ablation Controls

The following formulas contrast the information retained by alternative reliability maps computed from the same disagreements B_j, E_j:

$$\begin{aligned} \omega_{\mathrm{all}}(u) &= 1, & \omega_{\mathrm{bin}}(u) &= \mathbf{1}[E_j(u) < B_j(u)], \\ \omega_{\mathrm{abs}}(u) &= [B_j(u) - E_j(u)]_+, & \omega_{\mathrm{rel}}(u) &= \frac{[B_j(u) - E_j(u)]_+}{B_j(u) + \varepsilon}. \end{aligned} \tag{26}$$

Table 6: Information retained by the reliability alternatives. "Scale normalized" denotes invariance to a common positive rescaling of $B_j(u)$ and $E_j(u)$. For ω_{rel}, exact invariance assumes $\varepsilon = 0$ and $B_j(u) > 0$. With the stabilizer, it holds approximately when $B_j(u) \gg \varepsilon$.

Map	Improvement sign	Magnitude	Scale normalized
ω_{all}	No	No	Yes
ω_{bin}	Yes	No	Yes
ω_{abs}	Yes	Absolute	No
ω_{rel}	Yes	Relative	Yes

Uniform and binary maps discard improvement magnitude, whereas absolute reduction varies with residual scale; the relative score retains normalized improvement and gives zero weight when $B_j(u) = 0$. Table 3 separately compares active-pixel and fixed-HW reduction together with the K-fold sweep for both repair models on Mip-NeRF 360, Train 9, $N_r = 48$. Each comparison matches cached repairs, optimization, and evaluation views.

C.2.1 Post-hoc diagnostic metrics

Figure 4 uses ground truth only for analysis; it does not enter weight estimation or training. The diagnostic uses Difix3D at $N_r = 48$ across all Mip-NeRF 360 scenes for Train 3/6/9, using matched initial renders, repaired images, and cross-fitted weight maps. Let $q = (j, u)$ index the set $\mathcal{U}_s$ of repaired-view pixels for scene s, and write $W_q = \omega_j(u)$. The signed ground-truth error reduction produced by repair is

$$\Delta_q^{\mathrm{gt}} = \frac{1}{3}\sum_{c=1}^{3}\left(\left|S_j(u,c) - I_j^{\mathrm{gt}}(u,c)\right| - \left|F_j(u,c) - I_j^{\mathrm{gt}}(u,c)\right|\right), \tag{27}$$

so $\Delta_q^{\mathrm{gt}} > 0$ means that the repaired pixel is closer to ground truth. For any nonempty pixel subset $\mathcal{A} \subseteq \mathcal{U}_s$, define

$$P_s(\mathcal{A}) = \frac{100}{|\mathcal{A}|}\sum_{q\in\mathcal{A}} \mathbf{1}[\Delta_q^{\mathrm{gt}} > 0], \qquad D_s(\mathcal{A}) = \frac{100}{|\mathcal{A}|}\sum_{q\in\mathcal{A}} \Delta_q^{\mathrm{gt}}. \tag{28}$$

Panels (a,b) evaluate P_s and D_s on active pixels in weight intervals; panel (c) uses $0 < W_q < 0.3$ and $W_q \geq 0.7$; panel (d) uses five coarse intervals; and panel (e) evaluates the rejected set $\mathcal{Z}_s = \{q : W_q = 0\}$. The displayed D_s values are in units of 10^{-2} because RGB intensities are normalized to $[0, 1]$.

For the equal-mass control in panel (f), let

$$\overline{W}_s = \frac{1}{|\mathcal{U}_s|}\sum_{q\in\mathcal{U}_s} W_q, \quad G_s^{\mathrm{map}} = \frac{100}{|\mathcal{U}_s|}\sum_{q\in\mathcal{U}_s} W_q\Delta_q^{\mathrm{gt}}, \quad G_s^{\mathrm{uniform}} = \frac{100\overline{W}_s}{|\mathcal{U}_s|}\sum_{q\in\mathcal{U}_s}\Delta_q^{\mathrm{gt}}. \tag{29}$$

The two controls have identical total mass because $\sum_q W_q = \sum_q \overline{W}_s$. Every reported diagnostic is first computed within a scene and then averaged uniformly across scenes.

C.3 Full Quantitative Results

C.3.1 Difix3D on Mip-NeRF 360

The complete matrix covers Mip-NeRF 360 with $N_t \in \{3, 6, 9\}$, $N_r \in \{12, 24, 36, 48\}$, and $\lambda_g \in \{0.1, 0.2, \ldots, 1.0\}$. Repair-camera subsets are nested as N_r grows, and all branches within a cell share the sealed reconstruction, repaired images, camera poses, and continuation schedule.

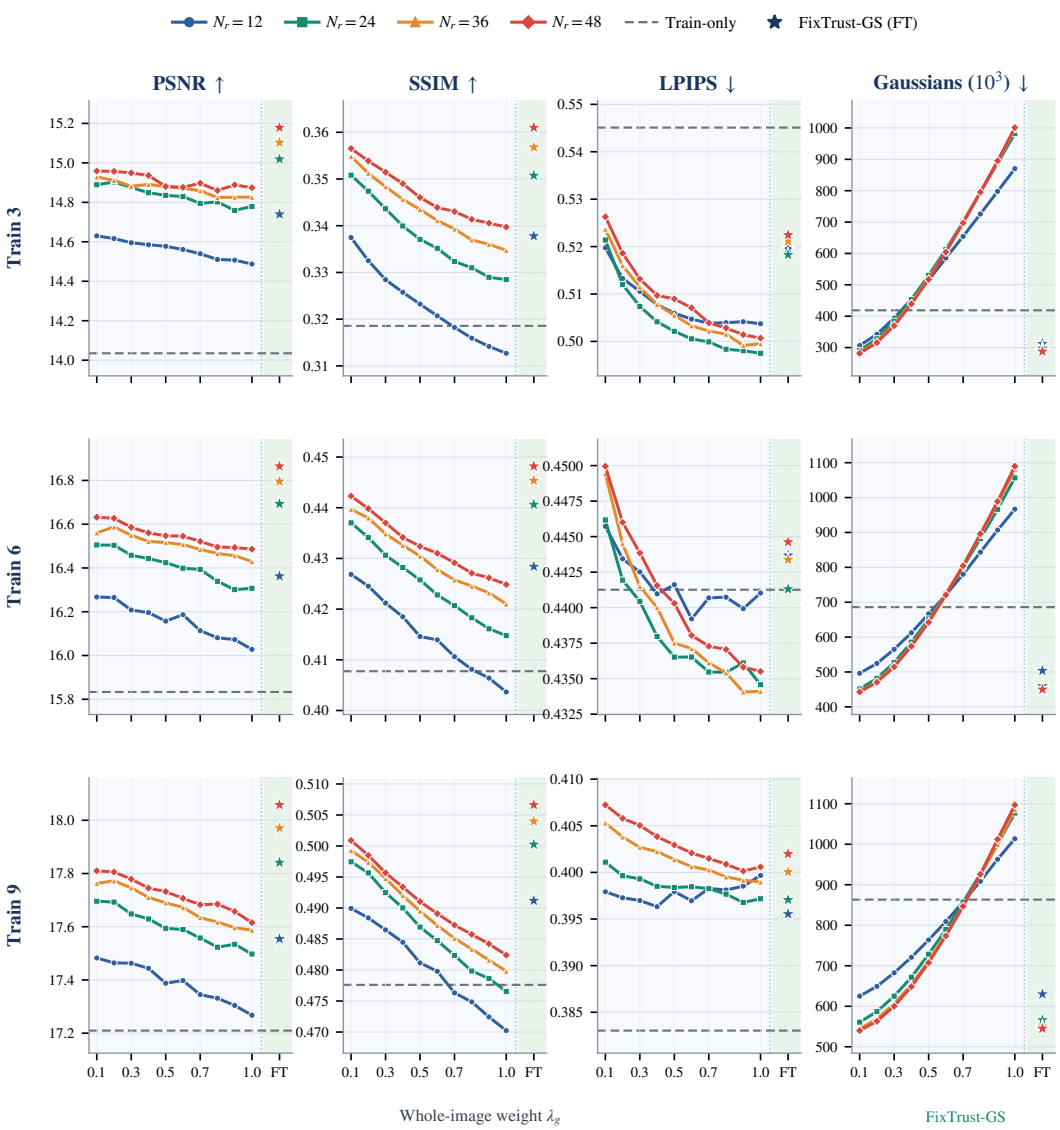

Figure 6: Complete Difix3D response on Mip-NeRF 360. Rows fix the captured training-view budget and colored curves fix N_r. Curves sweep the global weight $\lambda_g = 0.1$–1.0; stars denote FixTrust-GS, and the gray dashed line is Train-only. Values are scene-macro averages at step 8,999.

Result summary. At $N_r = 48$, FixTrust-GS reaches 15.18/16.86/18.06 dB for Train 3/6/9 while using 287.7/449.9/544.9 thousand Gaussians. It gives the strongest aggregate PSNR and SSIM throughout the three captured-view budgets, whereas the LPIPS ranking varies with supervision strength. The exact aggregate and scene-wise values below retain all repair budgets and the complete scalar-weight sweep; green and blue mark the best and second-best repaired-supervision results.

Table 7: Complete Difix3D weight sweep at Train 3, averaged equally over Mip-NeRF 360. The Train-only row is the paired $N_r = 48$ baseline.

Supervision	$N_r = 12$	$N_r = 24$	$N_r = 36$	$N_r = 48$
TRAIN-ONLY		14.04 / 0.319 0.545 / 418.8		
Global 1.0	14.49 / 0.313 **0.504** / 870.4	14.78 / 0.328 **0.497** / 982.2	14.83 / 0.335 0.500 / 996.0	14.87 / 0.340 **0.501** / 1001.1
Global 0.9	14.51 / 0.314 0.504 / 797.8	14.76 / 0.329 0.498 / 887.4	14.83 / 0.336 **0.499** / 891.5	14.89 / 0.341 0.501 / 895.2
Global 0.8	14.51 / 0.316 0.504 / 725.7	14.80 / 0.331 0.498 / 795.5	14.82 / 0.337 0.502 / 793.6	14.86 / 0.341 0.503 / 795.4
Global 0.7	14.54 / 0.318 0.504 / 654.3	14.79 / 0.332 0.500 / 700.8	14.86 / 0.339 0.502 / 699.6	14.90 / 0.343 0.504 / 697.2
Global 0.6	14.56 / 0.321 0.505 / 585.4	14.83 / 0.335 0.501 / 612.4	14.87 / 0.341 0.503 / 605.6	14.88 / 0.344 0.507 / 604.2
Global 0.5	14.58 / 0.323 0.506 / 514.8	14.84 / 0.337 0.502 / 530.4	14.88 / 0.343 0.506 / 521.9	14.88 / 0.346 0.509 / 517.2
Global 0.4	14.59 / 0.326 0.508 / 454.0	14.85 / 0.340 0.504 / 453.1	14.89 / 0.346 0.508 / 442.8	14.94 / 0.349 0.510 / 439.1
Global 0.3	14.60 / 0.328 0.511 / 394.1	14.88 / 0.344 0.507 / 384.0	14.88 / 0.348 0.511 / 374.5	14.95 / 0.351 0.513 / 369.7
Global 0.2	14.62 / 0.333 0.513 / 342.4	14.90 / 0.347 0.512 / 328.1	14.91 / 0.351 0.516 / 319.8	14.96 / 0.354 0.519 / 315.7
Global 0.1	14.63 / 0.337 0.520 / **306.5**	14.89 / **0.351** 0.521 / **290.8**	14.93 / 0.355 0.524 / **283.5**	14.96 / 0.357 0.526 / **281.8**
FIXTRUST-GS	**14.74** / **0.338** 0.520 / 312.2	**15.02** / 0.351 0.518 / 297.1	**15.10** / **0.357** 0.521 / 289.2	**15.18** / **0.361** 0.522 / 287.7

Table 8: Scene-wise Difix3D results at Train 3. Repair-view count varies for FIXTRUST-GS; Train-only is the paired $N_r = 48$ run.

Supervision	Metric	Bicycle	Bonsai	Counter	Flowers	Garden	Kitchen	Room	Stump	Treehill
TRAIN-ONLY	PSNR ↑	12.51	13.14	14.79	10.88	16.88	17.18	15.68	14.32	10.93
	SSIM ↑	0.218	0.343	0.373	0.170	0.362	0.510	0.494	0.175	0.222
	LPIPS ↓	0.690	0.536	0.459	0.642	0.473	0.341	0.436	0.634	0.696
	#GS ↓	324 (ref.)	335 (ref.)	193 (ref.)	644 (ref.)	495 (ref.)	297 (ref.)	166 (ref.)	800 (ref.)	516 (ref.)
FIXTRUST-GS $N_r = 12$	PSNR ↑	13.02	14.38	15.41	11.74	17.85	18.17	15.98	14.68	11.43
	SSIM ↑	0.225	0.369	0.388	0.194	0.398	0.545	0.486	0.188	0.249
	LPIPS ↓	0.649	0.496	0.434	**0.656**	**0.446**	0.303	0.411	**0.593**	**0.689**
	#GS ↓	278 (↓14.2%)	**309** (↓**7.9%**)	**174** (↓**9.7%**)	350 (↓45.6%)	343 (↓30.8%)	282 (↓5.0%)	152 (↓8.3%)	630 (↓21.2%)	292 (↓43.3%)
FIXTRUST-GS $N_r = 24$	PSNR ↑	**13.46**	14.75	15.62	12.01	17.99	18.16	16.39	**15.09**	11.69
	SSIM ↑	**0.242**	0.388	0.404	0.202	0.405	0.552	0.498	0.203	0.262
	LPIPS ↓	**0.641**	**0.489**	0.427	0.663	0.451	0.298	0.401	0.599	0.696
	#GS ↓	251 (↓22.3%)	309 (↓7.8%)	177 (↓8.5%)	313 (↓51.4%)	**342** (↓**30.9%**)	282 (↓4.9%)	149 (↓10.1%)	581 (↓27.3%)	269 (↓47.8%)
FIXTRUST-GS $N_r = 36$	PSNR ↑	13.46	14.76	15.74	12.20	18.06	18.25	16.58	14.96	11.92
	SSIM ↑	0.239	0.402	0.412	0.205	0.411	0.558	0.510	0.202	0.272
	LPIPS ↓	0.658	0.490	**0.427**	0.665	0.453	0.297	0.394	0.607	0.700
	#GS ↓	207 (↓36.1%)	311 (↓7.1%)	177 (↓8.5%)	**305** (↓**52.7%**)	343 (↓30.8%)	282 (↓4.8%)	149 (↓10.5%)	571 (↓28.6%)	259 (↓49.8%)
FIXTRUST-GS $N_r = 48$	PSNR ↑	13.38	**14.93**	**15.75**	**12.20**	**18.10**	**18.58**	**16.67**	14.94	**12.06**
	SSIM ↑	0.238	**0.408**	**0.414**	**0.209**	**0.414**	**0.566**	**0.519**	**0.205**	**0.277**
	LPIPS ↓	0.672	0.489	0.428	0.661	0.455	**0.294**	**0.389**	0.611	0.704
	#GS ↓	**200** (↓**38.1%**)	314 (↓6.3%)	178 (↓8.1%)	306 (↓52.4%)	343 (↓30.7%)	**280** (↓**5.5%**)	**146** (↓**12.4%**)	**564** (↓**29.5%**)	**258** (↓**50.0%**)

Train 3 result. At $N_r = 48$, FIXTRUST-GS improves PSNR by 1.14 dB and SSIM by 0.042 over Train-only with 31.3% fewer Gaussians; LPIPS decreases by 0.023. Macro PSNR for $N_r = 12/24/36/48$ is 14.74/15.02/15.10/15.18 dB. The scene-wise PSNR optimum spans $N_r = 24/48$, while mean Gaussian count changes from 312.2k to 287.7k.

Table 9: Complete Difix3D weight sweep at Train 6, averaged equally over Mip-NeRF 360. The Train-only row is the paired $N_r = 48$ baseline.

Supervision	$N_r = 12$	$N_r = 24$	$N_r = 36$	$N_r = 48$
TRAIN-ONLY		15.83 / 0.408 0.441 / 685.9		
Global 1.0	16.03 / 0.404 0.441 / 966.7	16.31 / 0.415 **0.435** / 1056.5	16.43 / 0.421 0.434 / 1081.3	16.49 / 0.425 **0.436** / 1089.8
Global 0.9	16.07 / 0.406 0.440 / 906.7	16.30 / 0.416 0.436 / 966.4	16.46 / 0.423 **0.434** / 983.5	16.49 / 0.426 0.436 / 988.9
Global 0.8	16.08 / 0.408 0.441 / 843.4	16.34 / 0.418 0.435 / 882.3	16.47 / 0.425 0.435 / 892.1	16.50 / 0.427 0.437 / 896.0
Global 0.7	16.11 / 0.411 0.441 / 780.1	16.39 / 0.421 0.435 / 801.3	16.48 / 0.426 0.436 / 805.7	16.52 / 0.429 0.437 / 804.5
Global 0.6	16.19 / 0.414 **0.439** / 720.7	16.40 / 0.423 0.437 / 724.0	16.51 / 0.428 0.437 / 722.9	16.55 / 0.431 0.438 / 720.7
Global 0.5	16.16 / 0.415 0.442 / 666.5	16.42 / 0.426 0.437 / 651.8	16.52 / 0.430 0.438 / 646.2	16.55 / 0.432 0.440 / 642.3
Global 0.4	16.20 / 0.418 0.441 / 612.3	16.44 / 0.428 0.438 / 585.4	16.52 / 0.433 0.440 / 575.1	16.56 / 0.434 0.442 / 573.7
Global 0.3	16.21 / 0.421 0.443 / 565.0	16.46 / 0.431 0.440 / 527.2	16.55 / 0.435 0.441 / 518.5	16.59 / 0.437 0.444 / 514.7
Global 0.2	16.26 / 0.425 0.443 / 524.5	16.50 / 0.434 0.442 / 481.8	16.59 / 0.438 0.444 / 472.3	16.63 / 0.440 0.446 / 470.4
Global 0.1	16.27 / 0.427 0.446 / **496.3**	16.51 / 0.437 0.446 / **451.4**	16.56 / 0.440 0.449 / **443.7**	16.63 / 0.442 0.450 / **442.8**
FIXTRUST-GS	**16.36 / 0.428** 0.444 / 503.6	**16.69 / 0.441** 0.441 / 459.7	**16.80 / 0.445** 0.443 / 451.2	**16.86 / 0.448** 0.445 / 449.9

Table 10: Scene-wise Difix3D results at Train 6. Repair-view count varies for FIXTRUST-GS; Train-only is the paired $N_r = 48$ run.

Supervision	Metric	Bicycle	Bonsai	Counter	Flowers	Garden	Kitchen	Room	Stump	Treehill
TRAIN-ONLY	PSNR ↑	13.79	15.42	16.74	11.98	19.27	19.76	17.37	15.96	12.21
	SSIM ↑	0.246	0.471	0.482	0.208	0.509	0.640	0.587	0.241	0.287
	LPIPS ↓	0.579	0.414	0.352	0.584	0.332	0.224	0.354	0.506	0.626
	#GS ↓	1132 (ref.)	521 (ref.)	275 (ref.)	965 (ref.)	726 (ref.)	364 (ref.)	267 (ref.)	1177 (ref.)	746 (ref.)
FIXTRUST-GS $N_r = 12$	PSNR ↑	14.44	15.87	17.23	12.57	19.96	20.17	17.76	16.66	12.60
	SSIM ↑	0.277	0.492	0.504	0.228	0.530	0.659	0.590	0.263	0.312
	LPIPS ↓	**0.579**	0.400	0.342	**0.612**	**0.340**	0.211	0.346	**0.527**	**0.637**
	#GS ↓	893 (↓21.2%)	462 (↓11.4%)	250 (↓9.1%)	603 (↓37.5%)	497 (↓31.6%)	**342 (↓5.9%)**	248 (↓6.9%)	815 (↓30.7%)	422 (↓43.4%)
FIXTRUST-GS $N_r = 24$	PSNR ↑	14.78	16.66	17.51	12.80	20.14	20.35	18.00	**17.17**	12.83
	SSIM ↑	0.291	0.522	0.516	0.234	0.536	0.667	0.600	0.278	0.321
	LPIPS ↓	0.583	0.378	0.333	0.620	0.346	0.207	0.334	0.527	0.645
	#GS ↓	690 (↓39.0%)	**460 (↓11.8%)**	248 (↓10.0%)	542 (↓43.8%)	**488 (↓32.7%)**	344 (↓5.5%)	235 (↓12.1%)	755 (↓35.9%)	376 (↓49.6%)
FIXTRUST-GS $N_r = 36$	PSNR ↑	**14.83**	16.63	**17.66**	13.03	20.25	20.41	18.15	17.12	13.10
	SSIM ↑	**0.292**	0.528	0.523	0.239	0.541	0.670	0.607	0.277	0.331
	LPIPS ↓	0.593	0.380	0.331	0.623	0.347	0.205	**0.328**	0.534	0.649
	#GS ↓	**657 (↓41.9%)**	462 (↓11.3%)	**245 (↓10.9%)**	530 (↓45.1%)	489 (↓32.6%)	347 (↓4.6%)	227 (↓14.9%)	738 (↓37.3%)	365 (↓51.1%)
FIXTRUST-GS $N_r = 48$	PSNR ↑	14.78	**16.81**	17.65	**13.13**	**20.27**	**20.50**	**18.20**	17.14	**13.30**
	SSIM ↑	0.289	**0.536**	**0.524**	**0.242**	**0.541**	**0.674**	**0.610**	**0.282**	**0.335**
	LPIPS ↓	0.600	**0.372**	**0.331**	0.622	0.351	**0.204**	0.330	0.537	0.655
	#GS ↓	663 (↓41.5%)	461 (↓11.5%)	246 (↓10.6%)	**526 (↓45.4%)**	490 (↓32.5%)	344 (↓5.3%)	**225 (↓15.8%)**	**730 (↓38.0%)**	**364 (↓51.2%)**

Train 6 result. At $N_r = 48$, FIXTRUST-GS improves PSNR by 1.03 dB and SSIM by 0.040 over Train-only with 34.4% fewer Gaussians; LPIPS increases by 0.003. Macro PSNR for $N_r = 12/24/36/48$ is 16.36/16.69/16.80/16.86 dB. The scene-wise PSNR optimum spans $N_r = 24/36/48$, while mean Gaussian count changes from 503.6k to 449.9k.

Table 11: Complete Difix3D weight sweep at Train 9, averaged equally over Mip-NeRF 360. The Train-only row is the paired $N_r = 48$ baseline.

Supervision	$N_r = 12$	$N_r = 24$	$N_r = 36$	$N_r = 48$
TRAIN-ONLY		17.21 / 0.478 0.383 / 863.3		
Global 1.0	17.27 / 0.470 0.400 / 1013.6	17.50 / 0.477 0.397 / 1077.0	17.59 / 0.480 **0.399** / 1086.1	17.62 / 0.482 0.401 / 1097.1
Global 0.9	17.30 / 0.472 0.399 / 962.8	17.53 / 0.479 **0.397** / 1000.1	17.60 / 0.482 0.399 / 999.5	17.66 / 0.484 **0.400** / 1012.2
Global 0.8	17.33 / 0.475 0.398 / 908.5	17.52 / 0.480 0.398 / 927.1	17.62 / 0.483 0.400 / 922.9	17.68 / 0.486 0.401 / 925.9
Global 0.7	17.34 / 0.476 0.398 / 857.9	17.56 / 0.482 0.398 / 857.2	17.63 / 0.485 0.400 / 847.5	17.68 / 0.487 0.401 / 846.9
Global 0.6	17.40 / 0.480 0.397 / 809.2	17.59 / 0.485 0.398 / 789.5	17.67 / 0.487 0.401 / 776.5	17.71 / 0.489 0.402 / 773.7
Global 0.5	17.39 / 0.481 0.398 / 763.7	17.59 / 0.487 0.398 / 728.6	17.69 / 0.489 0.401 / 712.1	17.73 / 0.491 0.403 / 707.6
Global 0.4	17.44 / 0.484 0.396 / 720.9	17.63 / 0.490 0.399 / 671.7	17.71 / 0.492 0.402 / 653.5	17.75 / 0.493 0.404 / 648.4
Global 0.3	17.46 / 0.486 0.397 / 682.6	17.65 / 0.492 0.399 / 625.0	17.75 / 0.495 0.403 / 606.3	17.78 / 0.496 0.405 / 600.2
Global 0.2	17.46 / 0.488 0.397 / 649.3	17.69 / 0.496 0.400 / 586.9	17.77 / 0.497 0.404 / 568.6	17.81 / 0.498 0.406 / 562.9
Global 0.1	17.48 / 0.490 0.398 / **625.0**	17.70 / 0.497 0.401 / **560.8**	17.76 / 0.499 0.405 / **544.7**	17.81 / 0.501 0.407 / **540.0**
FIXTRUST-GS	**17.55 / 0.491** **0.396** / 629.8	**17.84 / 0.500** 0.397 / 565.9	**17.97 / 0.504** 0.400 / 550.1	**18.06 / 0.507** 0.402 / 544.9

Table 12: Scene-wise Difix3D results at Train 9. Repair-view count varies for FIXTRUST-GS; Train-only is the paired $N_r = 48$ run.

Supervision	Metric	Bicycle	Bonsai	Counter	Flowers	Garden	Kitchen	Room	Stump	Treehill
TRAIN-ONLY	PSNR ↑	15.17	17.62	17.92	12.83	20.93	20.80	19.17	16.75	13.69
	SSIM ↑	0.305	0.598	0.555	0.248	0.611	0.719	0.649	0.273	0.342
	LPIPS ↓	0.527	0.311	0.290	0.554	0.266	0.161	0.293	0.474	0.571
	#GS ↓	1487 (ref.)	629 (ref.)	300 (ref.)	1150 (ref.)	802 (ref.)	473 (ref.)	370 (ref.)	1446 (ref.)	1111 (ref.)
FIXTRUST-GS $N_r = 12$	PSNR ↑	15.70	17.92	18.05	13.23	21.27	21.08	19.34	17.29	14.10
	SSIM ↑	0.328	0.608	0.567	0.264	0.611	0.731	0.659	0.293	0.360
	LPIPS ↓	**0.546**	0.304	0.293	**0.591**	**0.292**	0.154	0.290	**0.510**	**0.580**
	#GS ↓	1183 (↓20.4%)	543 (↓13.7%)	269 (↓10.3%)	731 (↓36.4%)	529 (↓34.1%)	422 (↓10.7%)	323 (↓12.7%)	1000 (↓30.8%)	666 (↓40.1%)
FIXTRUST-GS $N_r = 24$	PSNR ↑	16.07	18.30	18.35	13.45	21.48	21.37	19.56	17.69	14.29
	SSIM ↑	0.341	0.626	0.577	0.269	0.614	0.737	0.665	0.305	0.368
	LPIPS ↓	0.554	0.290	0.286	0.604	0.300	0.153	0.284	0.518	0.586
	#GS ↓	920 (↓38.1%)	**540 (↓14.2%)**	263 (↓12.4%)	650 (↓43.5%)	514 (↓35.9%)	422 (↓10.9%)	305 (↓17.7%)	895 (↓38.1%)	586 (↓47.3%)
FIXTRUST-GS $N_r = 36$	PSNR ↑	16.13	18.52	18.54	13.66	21.51	21.43	19.72	**17.70**	14.53
	SSIM ↑	0.342	0.633	0.584	0.273	**0.614**	0.739	0.669	0.307	0.375
	LPIPS ↓	0.564	0.287	0.284	0.605	0.307	0.153	0.282	0.527	0.593
	#GS ↓	859 (↓42.2%)	541 (↓14.0%)	**262 (↓12.6%)**	631 (↓45.1%)	**511 (↓36.4%)**	423 (↓10.7%)	293 (↓20.8%)	867 (↓40.0%)	562 (↓49.4%)
FIXTRUST-GS $N_r = 48$	PSNR ↑	**16.15**	**18.72**	**18.64**	**13.70**	**21.53**	**21.43**	**19.82**	17.68	**14.85**
	SSIM ↑	**0.343**	**0.639**	**0.586**	**0.274**	0.614	**0.741**	**0.673**	**0.309**	**0.382**
	LPIPS ↓	0.571	**0.281**	**0.283**	0.608	0.312	**0.152**	**0.279**	0.534	0.598
	#GS ↓	**847 (↓43.0%)**	542 (↓13.8%)	264 (↓12.2%)	**626 (↓45.5%)**	511 (↓36.3%)	**418 (↓11.6%)**	**291 (↓21.4%)**	**849 (↓41.3%)**	**556 (↓50.0%)**

Train 9 result. At $N_r = 48$, FIXTRUST-GS improves PSNR by 0.85 dB and SSIM by 0.029 over Train-only with 36.9% fewer Gaussians; LPIPS increases by 0.019. Macro PSNR for $N_r = 12/24/36/48$ is 17.55/17.84/17.97/18.06 dB. The scene-wise PSNR optimum spans $N_r = 36/48$, while mean Gaussian count changes from 629.8k to 544.9k.

C.3.2 DIFIX3D ON DEEP BLENDING

The complete matrix covers Deep Blending with $N_t \in \{3, 6, 9\}$, $N_r \in \{12, 24, 36, 48\}$, and $\lambda_g \in \{0.1, 0.2, \ldots, 1.0\}$. It follows the shared protocol in Appendix C.1, including nested repair-camera subsets and matched continuation branches within every evaluation cell.

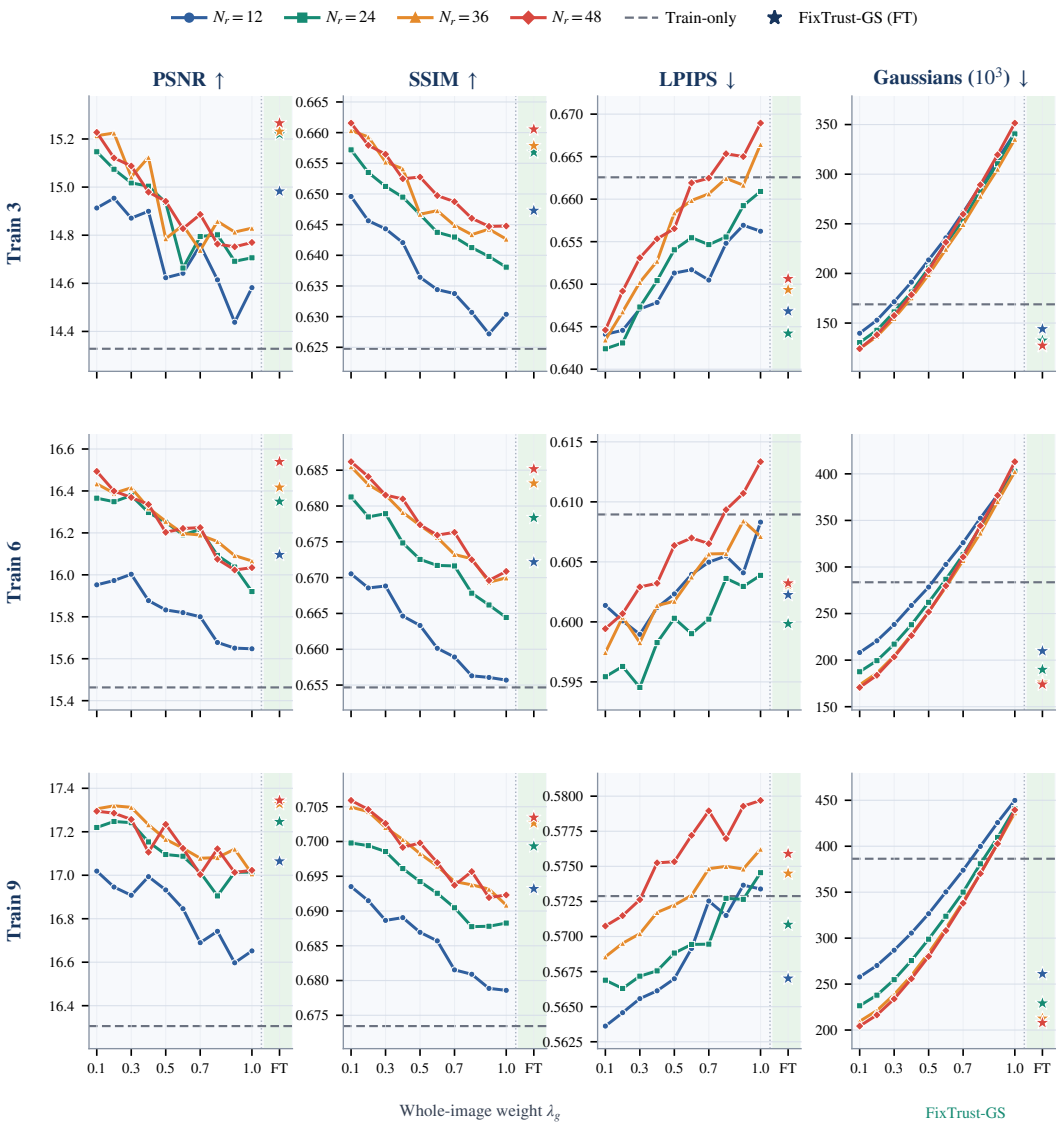

Figure 7: Complete Difix3D response on Deep Blending. Rows fix the captured training-view budget and colored curves fix N_r. Curves sweep the global weight $\lambda_g = 0.1$–1.0; stars denote FIXTRUST-GS, and the gray dashed line is Train-only. Values are scene-macro averages at step 8,999.

Result summary. At $N_r = 48$, FIXTRUST-GS reaches 15.27/16.54/17.34 dB for Train 3/6/9 while using 127.4/174.0/208.0 thousand Gaussians. The method improves aggregate PSNR and SSIM over Train-only throughout the captured-view budgets and avoids the Gaussian growth induced by large uniform weights. The tables below retain the complete repair-budget and scalar-weight sweeps together with scene-wise measurements.

Table 13: Complete Difix3D weight sweep at Train 3, averaged equally over Deep Blending. The Train-only row is the paired $N_r = 48$ baseline.

Supervision	$N_r = 12$	$N_r = 24$	$N_r = 36$	$N_r = 48$
TRAIN-ONLY			14.33 / 0.625 0.663 / 168.9	
Global 1.0	14.58 / 0.630 0.656 / 341.5	14.71 / 0.638 0.661 / 340.6	14.83 / 0.643 0.666 / 335.2	14.77 / 0.645 0.669 / 351.5
Global 0.9	14.44 / 0.627 0.657 / 315.1	14.69 / 0.640 0.659 / 310.5	14.81 / 0.644 0.662 / 305.1	14.75 / 0.645 0.665 / 319.5
Global 0.8	14.61 / 0.631 0.655 / 287.9	14.80 / 0.641 0.656 / 283.3	14.86 / 0.643 0.662 / 277.9	14.76 / 0.646 0.665 / 289.3
Global 0.7	14.76 / 0.634 0.650 / 261.5	14.79 / 0.643 0.655 / 253.7	14.74 / 0.645 0.661 / 249.4	14.89 / 0.649 0.662 / 259.8
Global 0.6	14.64 / 0.634 0.652 / 235.9	14.66 / 0.644 0.655 / 229.8	14.84 / 0.647 0.660 / 224.4	14.83 / 0.650 0.662 / 231.5
Global 0.5	14.62 / 0.636 0.651 / 213.6	14.94 / 0.647 0.654 / 205.3	14.79 / 0.647 0.658 / 198.8	14.94 / 0.653 0.657 / 203.0
Global 0.4	14.90 / 0.642 0.648 / 191.3	15.00 / 0.649 0.650 / 181.2	15.12 / 0.654 0.653 / 175.5	14.98 / 0.652 0.655 / 178.5
Global 0.3	14.87 / 0.644 0.647 / 171.5	15.02 / 0.651 0.647 / 161.3	15.04 / 0.655 0.650 / 154.7	15.09 / 0.656 0.653 / 157.5
Global 0.2	14.95 / 0.646 0.645 / 152.9	15.07 / 0.654 0.643 / 142.8	15.22 / 0.659 0.647 / 136.5	15.12 / 0.658 0.649 / 138.2
Global 0.1	14.91 / **0.650** **0.644** / **139.7**	15.15 / **0.657** **0.642** / **130.3**	15.21 / **0.660** **0.643** / **124.1**	15.23 / **0.662** **0.645** / **124.1**
FIXTRUST-GS	**14.98** / 0.647 0.647 / 144.1	**15.22** / 0.657 0.644 / 132.5	**15.23** / 0.658 0.649 / 126.9	**15.27** / 0.661 0.651 / 127.4

Table 14: Scene-wise Difix3D results at Train 3. Repair-view count varies for FIXTRUST-GS; Train-only is the paired $N_r = 48$ run.

Supervision	Metric	DrJohnson	Playroom	Bedroom	Boats	Bridge	CreepyAttic
	PSNR ↑	16.23	13.48	14.67	10.36	14.99	16.24
	SSIM ↑	0.652	0.641	0.636	0.664	0.532	0.623
	LPIPS ↓	0.671	0.650	0.610	0.599	0.664	0.780
	#GS ↓	155 (ref.)	126 (ref.)	124 (ref.)	236 (ref.)	127 (ref.)	246 (ref.)
FIXTRUST-GS $N_r = 12$	PSNR ↑	17.20	14.97	14.77	12.12	15.07	**15.77**
	SSIM ↑	0.677	0.682	0.635	0.730	0.548	0.612
	LPIPS ↓	0.649	0.626	0.605	0.575	0.680	**0.745**
	#GS ↓	126 (↓18.9%)	119 (↓5.9%)	**128** (↑3.4%)	188 (↓20.2%)	104 (↓18.1%)	200 (↓18.7%)
FIXTRUST-GS $N_r = 24$	PSNR ↑	17.14	15.16	**15.06**	12.67	15.57	15.73
	SSIM ↑	0.678	0.690	**0.650**	0.749	0.561	0.613
	LPIPS ↓	0.646	0.624	**0.596**	**0.564**	**0.677**	0.759
	#GS ↓	125 (↓19.1%)	116 (↓8.1%)	103 (↓17.2%)	184 (↓22.0%)	93 (↓26.9%)	175 (↓28.8%)
FIXTRUST-GS $N_r = 36$	PSNR ↑	17.15	15.57	14.74	12.58	**15.64**	15.71
	SSIM ↑	0.679	0.701	0.642	0.747	**0.563**	**0.615**
	LPIPS ↓	**0.644**	0.619	0.614	0.569	0.678	0.772
	#GS ↓	**124** (↓20.1%)	**110** (↓12.5%)	100 (↓19.2%)	**172** (↓27.0%)	92 (↓27.8%)	**164** (↓33.4%)
FIXTRUST-GS $N_r = 48$	PSNR ↑	**17.28**	**15.87**	14.82	**12.69**	15.48	15.45
	SSIM ↑	**0.685**	**0.704**	0.648	**0.752**	0.563	0.612
	LPIPS ↓	0.646	**0.617**	0.600	0.568	0.687	0.786
	#GS ↓	124 (↓20.0%)	112 (↓11.1%)	**97** (↓21.8%)	178 (↓24.6%)	**90** (↓29.2%)	164 (↓33.4%)

Train 3 result. At $N_r = 48$, FIXTRUST-GS changes PSNR by +0.94 dB and SSIM by +0.036 over Train-only with 24.6% fewer Gaussians; LPIPS decreases by 0.012. Macro PSNR for $N_r = 12/24/36/48$ is 14.98/15.22/15.23/15.27 dB. The scene-wise PSNR optimum spans $N_r = 12/24/36/48$, while mean Gaussian count changes from 144.1k to 127.4k.

Table 15: Complete Difix3D weight sweep at Train 6, averaged equally over Deep Blending. The Train-only row is the paired $N_r = 48$ baseline.

Supervision	$N_r = 12$	$N_r = 24$	$N_r = 36$	$N_r = 48$
TRAIN-ONLY		15.46 / 0.655 0.609 / 283.6		
Global 1.0	15.65 / 0.656 0.608 / 403.7	15.92 / 0.664 0.604 / 402.8	16.07 / 0.670 0.607 / 402.4	16.03 / 0.671 0.613 / 413.0
Global 0.9	15.65 / 0.656 0.604 / 377.4	16.04 / 0.666 0.603 / 373.2	16.09 / 0.669 0.608 / 370.2	16.02 / 0.670 0.611 / 376.8
Global 0.8	15.68 / 0.656 0.605 / 352.3	16.09 / 0.668 0.604 / 342.5	16.16 / 0.673 0.606 / 336.4	16.07 / 0.672 0.609 / 344.3
Global 0.7	15.80 / 0.659 0.605 / 326.1	16.22 / 0.672 0.600 / 313.1	16.19 / 0.673 0.606 / 307.1	16.23 / 0.676 0.607 / 310.8
Global 0.6	15.82 / 0.660 0.604 / 302.7	16.19 / 0.672 0.599 / 286.8	16.20 / 0.676 0.604 / 278.5	16.22 / 0.676 0.607 / 279.7
Global 0.5	15.83 / 0.663 0.602 / 278.5	16.25 / 0.673 0.600 / 262.1	16.26 / 0.677 0.602 / 251.8	16.20 / 0.677 0.606 / 251.7
Global 0.4	15.88 / 0.665 0.601 / 258.7	16.30 / 0.675 0.598 / 238.0	16.32 / 0.679 0.601 / 227.4	16.34 / 0.681 0.603 / 226.3
Global 0.3	16.00 / 0.669 **0.599** / 238.3	**16.38** / 0.679 **0.595** / 217.2	16.41 / 0.682 0.598 / 204.5	16.37 / 0.682 0.603 / 203.4
Global 0.2	15.97 / 0.669 0.600 / 220.8	16.35 / 0.678 0.596 / 199.4	16.39 / 0.683 0.600 / 186.1	16.40 / 0.684 0.601 / 183.8
Global 0.1	15.95 / 0.671 0.601 / **208.2**	16.37 / **0.681** 0.595 / **187.5**	**16.43** / **0.685** **0.597** / **173.5**	16.49 / **0.686** **0.599** / **170.7**
FIXTRUST-GS	**16.10** / **0.672** 0.602 / 209.8	16.35 / 0.678 0.600 / 189.8	16.42 / 0.683 0.603 / 176.1	**16.54** / 0.685 0.603 / 174.0

Table 16: Scene-wise Difix3D results at Train 6. Repair-view count varies for FIXTRUST-GS; Train-only is the paired $N_r = 48$ run.

Supervision	Metric	DrJohnson	Playroom	Bedroom	Boats	Bridge	CreepyAttic
	PSNR ↑	16.22	15.68	14.58	14.72	15.01	16.57
	SSIM ↑	0.657	0.690	0.638	0.769	0.542	0.632
	LPIPS ↓	0.659	0.550	0.609	0.457	0.666	0.712
	#GS ↓	259 (ref.)	225 (ref.)	209 (ref.)	422 (ref.)	173 (ref.)	413 (ref.)
FIXTRUST-GS $N_r = 12$	PSNR ↑	16.90	17.09	15.30	15.27	15.45	16.57
	SSIM ↑	0.677	0.721	0.657	0.785	0.562	0.630
	LPIPS ↓	0.645	0.538	0.596	0.464	0.685	**0.685**
	#GS ↓	177 (↓31.6%)	180 (↓19.7%)	158 (↓24.7%)	347 (↓17.9%)	110 (↓36.6%)	287 (↓30.5%)
FIXTRUST-GS $N_r = 24$	PSNR ↑	17.11	17.36	15.35	15.92	**15.89**	16.48
	SSIM ↑	0.684	0.729	0.663	0.798	0.570	0.625
	LPIPS ↓	**0.639**	0.538	0.584	**0.460**	**0.685**	0.693
	#GS ↓	170 (↓34.4%)	167 (↓25.6%)	136 (↓35.1%)	325 (↓23.1%)	96 (↓44.6%)	245 (↓40.7%)
FIXTRUST-GS $N_r = 36$	PSNR ↑	17.03	17.39	15.84	15.94	15.76	16.55
	SSIM ↑	0.684	0.732	**0.675**	0.798	**0.574**	0.636
	LPIPS ↓	0.647	0.539	**0.577**	0.465	0.689	0.701
	#GS ↓	**166** (↓36.0%)	158 (↓29.8%)	134 (↓36.1%)	282 (↓33.3%)	93 (↓46.2%)	224 (↓45.7%)
FIXTRUST-GS $N_r = 48$	PSNR ↑	**17.32**	**17.59**	**15.94**	**16.04**	15.71	**16.64**
	SSIM ↑	**0.691**	**0.736**	0.673	**0.801**	0.572	**0.638**
	LPIPS ↓	0.641	**0.532**	0.582	0.466	0.693	0.705
	#GS ↓	167 (↓35.5%)	**156** (↓30.4%)	**130** (↓37.8%)	**276** (↓34.7%)	**92** (↓46.9%)	**222** (↓46.2%)

Train 6 result. At $N_r = 48$, FIXTRUST-GS changes PSNR by +1.08 dB and SSIM by +0.031 over Train-only with 38.6% fewer Gaussians; LPIPS decreases by 0.006. Macro PSNR for $N_r = 12/24/36/48$ is 16.10/16.35/16.42/16.54 dB. The scene-wise PSNR optimum spans $N_r = 24/48$, while mean Gaussian count changes from 209.8k to 174.0k.

Table 17: Complete Difix3D weight sweep at Train 9, averaged equally over Deep Blending. The Train-only row is the paired $N_r = 48$ baseline.

Supervision	$N_r = 12$	$N_r = 24$	$N_r = 36$	$N_r = 48$
TRAIN-ONLY			16.31 / 0.673 0.573 / 386.4	
Global 1.0	16.65 / 0.679 0.573 / 449.9	17.01 / 0.688 0.575 / 440.9	17.01 / 0.691 0.576 / 436.5	17.02 / 0.692 0.580 / 439.5
Global 0.9	16.60 / 0.679 0.574 / 425.7	17.01 / 0.688 0.573 / 409.6	17.12 / 0.693 0.575 / 402.7	17.01 / 0.692 0.579 / 402.7
Global 0.8	16.74 / 0.681 0.571 / 399.8	16.91 / 0.688 0.573 / 380.9	17.08 / 0.694 0.575 / 370.1	17.12 / 0.696 0.577 / 370.2
Global 0.7	16.69 / 0.682 0.573 / 373.9	17.01 / 0.690 0.569 / 350.1	17.08 / 0.694 0.575 / 339.8	17.00 / 0.694 0.579 / 338.0
Global 0.6	16.85 / 0.686 0.569 / 350.3	17.09 / 0.693 0.569 / 323.8	17.12 / 0.696 0.573 / 310.7	17.12 / 0.697 0.577 / 308.2
Global 0.5	16.93 / 0.687 0.567 / 326.6	17.10 / 0.694 0.569 / 298.9	17.17 / 0.698 0.572 / 283.8	17.23 / 0.700 0.575 / 280.1
Global 0.4	16.99 / 0.689 0.566 / 305.5	17.15 / 0.696 0.568 / 275.6	17.23 / 0.700 0.572 / 259.3	17.11 / 0.699 0.575 / 255.8
Global 0.3	16.91 / 0.689 0.566 / 286.9	17.24 / 0.699 0.567 / 254.8	17.31 / 0.702 0.570 / 238.7	17.26 / 0.703 0.573 / 233.8
Global 0.2	16.95 / 0.691 0.565 / 270.2	**17.25** / 0.699 **0.566** / 238.0	17.32 / 0.704 0.570 / 221.2	17.29 / 0.705 0.571 / 216.4
Global 0.1	17.02 / **0.694** **0.564** / **257.8**	17.22 / **0.700** 0.567 / **226.4**	17.31 / **0.705** **0.569** / **209.5**	17.29 / **0.706** **0.571** / **204.3**
FIXTRUST-GS	**17.06** / 0.693 0.567 / 261.1	17.25 / 0.699 0.571 / 229.2	**17.33** / 0.703 0.574 / 212.7	**17.34** / 0.703 0.576 / 208.0

Table 18: Scene-wise Difix3D results at Train 9. Repair-view count varies for FIXTRUST-GS; Train-only is the paired $N_r = 48$ run.

Supervision	Metric	DrJohnson	Playroom	Bedroom	Boats	Bridge	CreepyAttic
	PSNR ↑	16.55	16.02	15.61	16.48	15.75	17.44
	SSIM ↑	0.661	0.698	0.670	0.795	0.571	0.644
	LPIPS ↓	0.648	0.530	0.568	0.420	0.636	0.635
	#GS ↓	393 (ref.)	316 (ref.)	273 (ref.)	516 (ref.)	173 (ref.)	647 (ref.)
FIXTRUST-GS $N_r = 12$	PSNR ↑	17.39	17.27	16.32	17.02	16.33	18.06
	SSIM ↑	0.681	0.727	0.692	0.809	0.595	0.655
	LPIPS ↓	0.636	**0.523**	0.544	**0.433**	**0.652**	**0.614**
	#GS ↓	231 (↓41.3%)	229 (↓27.5%)	203 (↓25.7%)	402 (↓22.2%)	115 (↓33.4%)	386 (↓40.2%)
FIXTRUST-GS $N_r = 24$	PSNR ↑	17.81	17.59	**16.56**	17.13	16.37	18.02
	SSIM ↑	0.691	0.733	0.699	0.813	0.602	0.658
	LPIPS ↓	**0.634**	0.527	0.544	0.439	0.660	0.620
	#GS ↓	202 (↓48.6%)	208 (↓34.4%)	178 (↓34.8%)	371 (↓28.1%)	102 (↓41.0%)	314 (↓51.4%)
FIXTRUST-GS $N_r = 36$	PSNR ↑	17.90	**17.80**	16.52	17.13	16.58	18.04
	SSIM ↑	0.694	**0.740**	0.699	0.814	0.609	0.661
	LPIPS ↓	0.636	0.528	0.547	0.445	0.667	0.624
	#GS ↓	195 (↓50.3%)	198 (↓37.3%)	174 (↓36.3%)	333 (↓35.4%)	99 (↓42.6%)	276 (↓57.3%)
FIXTRUST-GS $N_r = 48$	PSNR ↑	**17.94**	17.74	16.44	**17.17**	**16.62**	**18.16**
	SSIM ↑	**0.696**	0.738	**0.699**	**0.815**	**0.609**	**0.662**
	LPIPS ↓	0.640	0.530	**0.543**	0.448	0.667	0.628
	#GS ↓	**190** (↓51.6%)	**196** (↓37.9%)	**172** (↓37.2%)	**323** (↓37.5%)	**98** (↓43.6%)	**269** (↓58.4%)

Train 9 result. At $N_r = 48$, FIXTRUST-GS changes PSNR by +1.04 dB and SSIM by +0.030 over Train-only with 46.2% fewer Gaussians; LPIPS increases by 0.003. Macro PSNR for $N_r = 12/24/36/48$ is 17.06/17.25/17.33/17.34 dB. The scene-wise PSNR optimum spans $N_r = 24/36/48$, while mean Gaussian count changes from 261.1k to 208.0k.

C.3.3 Diffusion Harmonizer on Mip-NeRF 360

The complete matrix covers Mip-NeRF 360 with $N_t \in \{3, 6, 9\}$, $N_r \in \{12, 24, 36, 48\}$, and $\lambda_g \in \{0.1, 0.2, \ldots, 1.0\}$. It follows the shared protocol in Appendix C.1, including nested repair-camera subsets and matched continuation branches within every evaluation cell.

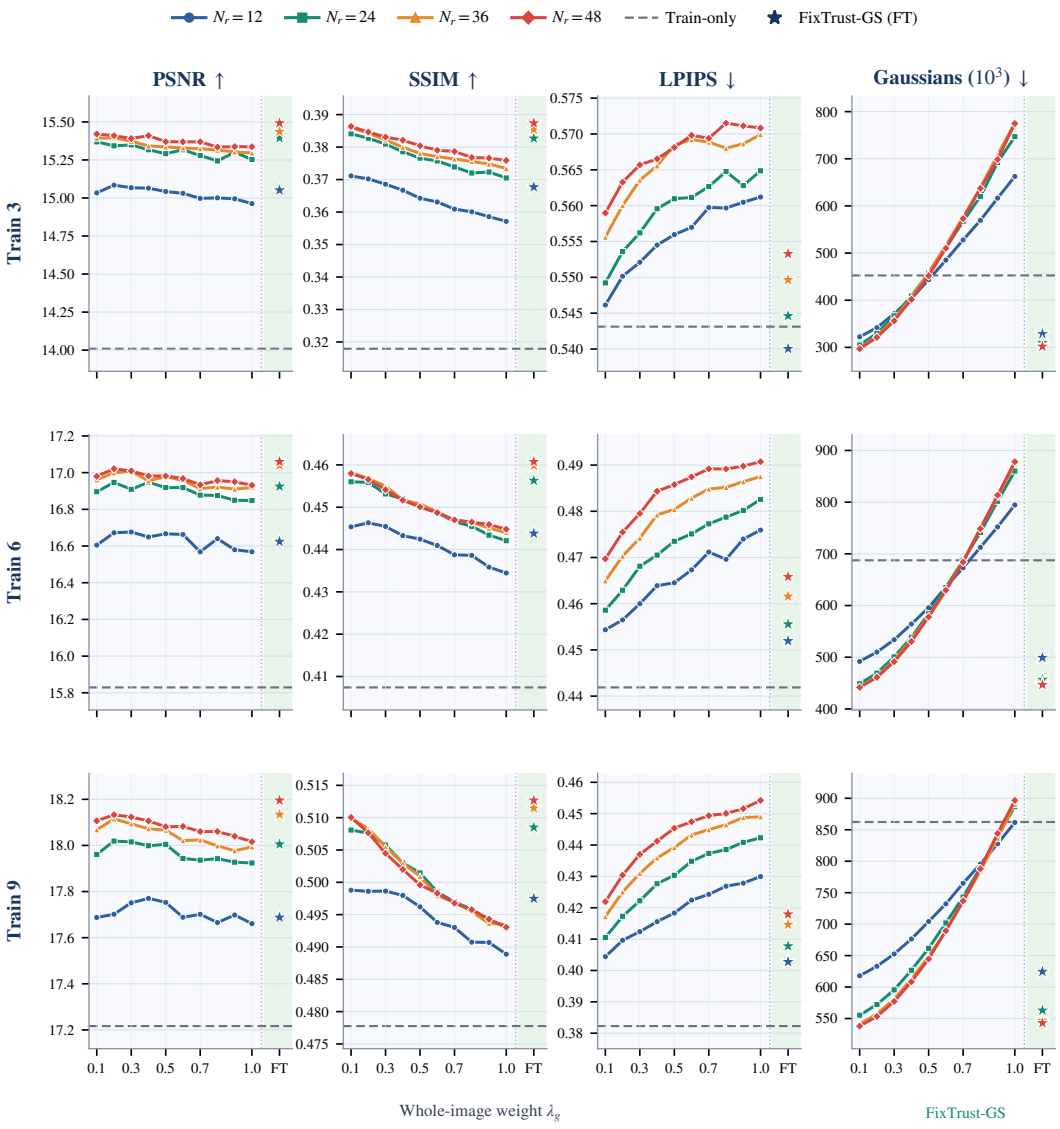

Figure 8: Complete Diffusion Harmonizer response on Mip-NeRF 360. Rows fix the captured training-view budget and colored curves fix N_r. Curves sweep the global weight $\lambda_g = 0.1$–1.0; stars denote FixTrust-GS, and the gray dashed line is Train-only. Values are scene-macro averages at step 8,999.

Result summary. At $N_r = 48$, FixTrust-GS reaches 15.49/17.06/18.20 dB for Train 3/6/9 while using 301.9/446.8/542.9 thousand Gaussians. Across the three captured-view budgets, it provides the strongest aggregate PSNR–SSIM balance among repaired supervision variants and substantially reduces Gaussian count relative to Global 1.0. Complete aggregate and scene-wise measurements follow.

Table 19: Complete Diffusion Harmonizer weight sweep at Train 3, averaged equally over Mip-NeRF 360. The Train-only row is the paired $N_r = 48$ baseline.

Supervision	$N_r = 12$	$N_r = 24$	$N_r = 36$	$N_r = 48$
TRAIN-ONLY		14.01 / 0.318 0.543 / 452.6		
Global 1.0	14.96 / 0.357 0.561 / 662.5	15.25 / 0.370 0.565 / 747.0	15.30 / 0.373 0.570 / 775.7	15.34 / 0.376 0.571 / 774.8
Global 0.9	14.99 / 0.359 0.560 / 616.5	15.30 / 0.372 0.563 / 692.2	15.30 / 0.375 0.569 / 708.3	15.34 / 0.377 0.571 / 698.5
Global 0.8	15.00 / 0.360 0.560 / 569.3	15.24 / 0.372 0.565 / 619.8	15.31 / 0.376 0.568 / 638.1	15.33 / 0.377 0.572 / 637.3
Global 0.7	15.00 / 0.361 0.560 / 527.8	15.28 / 0.374 0.563 / 567.4	15.32 / 0.376 0.569 / 576.3	15.37 / 0.379 0.569 / 573.1
Global 0.6	15.03 / 0.363 0.557 / 485.1	15.32 / 0.376 0.561 / 511.1	15.33 / 0.377 0.569 / 516.2	15.37 / 0.379 0.570 / 510.2
Global 0.5	15.04 / 0.364 0.556 / 443.5	15.29 / 0.377 0.561 / 457.3	15.34 / 0.378 0.568 / 458.9	15.37 / 0.380 0.568 / 451.9
Global 0.4	15.06 / 0.367 0.555 / 403.7	15.32 / 0.379 0.560 / 408.8	15.34 / 0.380 0.566 / 405.7	15.41 / 0.382 0.567 / 401.9
Global 0.3	15.07 / 0.369 0.552 / 371.5	15.35 / 0.381 0.556 / 366.9	15.37 / 0.382 0.564 / 360.4	15.39 / 0.383 0.566 / 356.2
Global 0.2	**15.08** / 0.370 0.550 / 341.8	15.34 / 0.383 0.554 / 329.6	15.40 / 0.384 0.560 / 323.5	15.41 / 0.385 0.563 / 321.3
Global 0.1	15.03 / **0.371** 0.546 / **322.4**	15.37 / **0.384** 0.549 / **305.2**	15.39 / **0.386** 0.556 / **298.6**	15.42 / 0.386 0.559 / **297.1**
FIXTRUST-GS	15.05 / 0.368 **0.540** / 328.5	**15.39** / 0.383 **0.545** / 310.0	**15.44** / 0.385 **0.550** / 303.4	**15.49** / **0.387** **0.553** / 301.9

Table 20: Scene-wise Diffusion Harmonizer results at Train 3. Repair-view count varies for FIXTRUST-GS; Train-only is the paired $N_r = 48$ run.

Supervision	Metric	Bicycle	Bonsai	Counter	Flowers	Garden	Kitchen	Room	Stump	Treehill
	PSNR ↑	12.33	13.17	14.80	10.87	16.86	17.09	15.70	14.32	10.94
	SSIM ↑	0.217	0.343	0.372	0.169	0.362	0.506	0.495	0.175	0.222
	LPIPS ↓	0.673	0.535	0.459	0.641	0.472	0.343	0.436	0.635	0.693
	#GS ↓	646 (ref.)	334 (ref.)	192 (ref.)	630 (ref.)	497 (ref.)	296 (ref.)	166 (ref.)	797 (ref.)	515 (ref.)
FIXTRUST-GS $N_r = 12$	PSNR ↑	13.31	14.38	15.90	11.96	18.49	18.70	16.23	14.96	11.53
	SSIM ↑	0.255	0.403	0.427	0.211	0.422	0.577	0.529	0.213	0.272
	LPIPS ↓	**0.676**	0.515	**0.459**	**0.692**	**0.471**	0.282	0.413	**0.642**	**0.709**
	#GS ↓	440 (↓31.9%)	305 (↓8.9%)	**177 (↓7.9%)**	349 (↓44.6%)	346 (↓30.5%)	283 (↓4.6%)	153 (↓8.0%)	626 (↓21.4%)	279 (↓45.8%)
FIXTRUST-GS $N_r = 24$	PSNR ↑	13.78	14.60	16.03	12.36	18.68	18.99	16.82	**15.31**	11.97
	SSIM ↑	**0.273**	0.426	0.441	0.219	0.430	0.588	0.552	0.226	0.289
	LPIPS ↓	0.684	0.518	0.465	0.704	0.479	0.275	0.411	0.651	0.715
	#GS ↓	386 (↓40.2%)	**304 (↓9.0%)**	178 (↓7.2%)	310 (↓50.7%)	**344 (↓30.9%)**	283 (↓4.6%)	147 (↓11.2%)	578 (↓27.4%)	260 (↓49.6%)
FIXTRUST-GS $N_r = 36$	PSNR ↑	**13.82**	14.62	**16.13**	12.39	**18.79**	18.94	16.86	15.27	12.11
	SSIM ↑	0.269	0.430	**0.448**	0.219	**0.433**	0.589	0.558	0.228	0.296
	LPIPS ↓	0.696	0.516	0.471	0.715	0.486	**0.273**	0.408	0.662	0.720
	#GS ↓	347 (↓46.2%)	305 (↓8.8%)	179 (↓6.8%)	303 (↓51.8%)	345 (↓30.7%)	283 (↓4.5%)	147 (↓11.4%)	568 (↓28.7%)	254 (↓50.7%)
FIXTRUST-GS $N_r = 48$	PSNR ↑	13.77	**14.79**	16.12	**12.49**	18.78	**19.07**	**16.93**	15.27	**12.22**
	SSIM ↑	0.266	**0.436**	0.447	**0.222**	0.433	**0.591**	**0.561**	**0.231**	**0.300**
	LPIPS ↓	0.709	**0.515**	0.476	0.712	0.491	0.274	**0.405**	0.671	0.726
	#GS ↓	**344 (↓46.8%)**	306 (↓8.5%)	180 (↓6.3%)	**303 (↓51.9%)**	345 (↓30.5%)	**281 (↓5.0%)**	**145 (↓12.4%)**	**562 (↓29.5%)**	**251 (↓51.3%)**

Train 3 result. At $N_r = 48$, FIXTRUST-GS changes PSNR by +1.48 dB and SSIM by +0.069 over Train-only with 33.3% fewer Gaussians; LPIPS increases by 0.010. Macro PSNR for $N_r = 12/24/36/48$ is 15.05/15.39/15.44/15.49 dB. The scene-wise PSNR optimum spans $N_r = 24/36/48$, while mean Gaussian count changes from 328.5k to 301.9k.

Table 21: Complete Diffusion Harmonizer weight sweep at Train 6, averaged equally over Mip-NeRF 360. The Train-only row is the paired $N_r = 48$ baseline.

Supervision	$N_r = 12$	$N_r = 24$	$N_r = 36$	$N_r = 48$
TRAIN-ONLY		15.83 / 0.407 0.442 / 687.5		
Global 1.0	16.57 / 0.434 0.476 / 794.6	16.85 / 0.442 0.483 / 860.2	16.92 / 0.444 0.487 / 874.1	16.93 / 0.445 0.491 / 878.2
Global 0.9	16.58 / 0.436 0.474 / 752.1	16.85 / 0.443 0.480 / 801.8	16.91 / 0.445 0.486 / 809.9	16.95 / 0.446 0.490 / 813.6
Global 0.8	16.64 / 0.439 0.470 / 712.5	16.88 / 0.445 0.479 / 741.2	16.92 / 0.446 0.485 / 749.4	16.96 / 0.446 0.489 / 748.6
Global 0.7	16.57 / 0.439 0.471 / 673.0	16.88 / 0.447 0.477 / 687.4	16.91 / 0.447 0.485 / 686.0	16.93 / 0.447 0.489 / 684.0
Global 0.6	16.66 / 0.441 0.467 / 636.0	16.92 / 0.449 0.475 / 634.9	16.96 / 0.449 0.483 / 629.7	16.97 / 0.449 0.487 / 629.7
Global 0.5	16.67 / 0.442 0.465 / 595.8	16.92 / 0.450 0.474 / 584.1	16.98 / 0.451 0.480 / 579.5	16.98 / 0.450 0.486 / 577.8
Global 0.4	16.65 / 0.443 0.464 / 564.2	**16.95** / 0.452 0.471 / 538.4	16.96 / 0.452 0.479 / 534.3	16.98 / 0.452 0.484 / 530.4
Global 0.3	**16.68** / 0.445 0.460 / 533.8	16.91 / 0.453 0.468 / 500.6	17.01 / 0.455 0.474 / 493.8	17.01 / 0.454 0.480 / 491.3
Global 0.2	16.67 / **0.446** 0.456 / 510.0	16.95 / 0.456 0.463 / 469.2	17.00 / 0.457 0.470 / 462.4	17.02 / 0.457 0.475 / 461.1
Global 0.1	16.60 / 0.445 0.454 / **491.9**	16.90 / 0.456 0.459 / **448.7**	16.96 / 0.458 0.465 / **441.8**	16.98 / 0.458 0.470 / **441.6**
FIXTRUST-GS	16.62 / 0.444 **0.452** / 499.2	16.93 / **0.456** **0.456** / 456.9	**17.04** / **0.460** **0.462** / 448.0	**17.06** / **0.461** **0.466** / 446.8

Table 22: Scene-wise Diffusion Harmonizer results at Train 6. Repair-view count varies for FIXTRUST-GS; Train-only is the paired $N_r = 48$ run.

Supervision	Metric	Bicycle	Bonsai	Counter	Flowers	Garden	Kitchen	Room	Stump	Treehill
	PSNR ↑	13.82	15.44	16.75	11.97	19.26	19.80	17.29	15.98	12.14
	SSIM ↑	0.247	0.470	0.481	0.207	0.509	0.641	0.583	0.242	0.286
	LPIPS ↓	0.579	0.414	0.352	0.586	0.332	0.223	0.357	0.506	0.628
	#GS ↓	1133 (ref.)	519 (ref.)	274 (ref.)	970 (ref.)	730 (ref.)	365 (ref.)	268 (ref.)	1180 (ref.)	749 (ref.)
FIXTRUST-GS $N_r = 12$	PSNR ↑	14.67	16.14	17.55	12.89	20.37	20.32	17.77	17.26	12.64
	SSIM ↑	0.293	0.510	0.522	0.241	0.540	0.673	0.610	0.281	0.323
	LPIPS ↓	**0.594**	0.408	0.344	**0.630**	**0.342**	0.198	0.344	**0.561**	**0.647**
	#GS ↓	883 (↓22.1%)	**459 (↓11.6%)**	252 (↓8.0%)	583 (↓39.9%)	498 (↓31.7%)	**341 (↓6.6%)**	239 (↓10.7%)	826 (↓30.0%)	410 (↓45.2%)
FIXTRUST-GS $N_r = 24$	PSNR ↑	14.97	16.46	17.79	13.11	20.62	20.67	17.89	17.86	12.95
	SSIM ↑	**0.306**	0.526	0.537	0.248	0.546	0.687	0.622	0.300	0.335
	LPIPS ↓	0.602	0.401	**0.340**	0.644	0.349	0.188	**0.342**	0.580	0.655
	#GS ↓	682 (↓39.9%)	459 (↓11.5%)	251 (↓8.6%)	527 (↓45.7%)	**489 (↓32.9%)**	345 (↓5.7%)	227 (↓15.3%)	764 (↓35.2%)	368 (↓50.8%)
FIXTRUST-GS $N_r = 36$	PSNR ↑	**14.98**	**16.55**	17.93	13.40	20.74	20.78	17.91	**17.88**	13.19
	SSIM ↑	0.303	**0.534**	0.543	0.253	**0.549**	0.691	0.625	0.300	0.342
	LPIPS ↓	0.616	**0.398**	0.341	0.651	0.355	**0.188**	0.344	0.602	0.660
	#GS ↓	**649 (↓42.8%)**	459 (↓11.5%)	**248 (↓9.5%)**	**514 (↓47.0%)**	491 (↓32.7%)	345 (↓5.4%)	224 (↓16.5%)	745 (↓36.9%)	357 (↓52.3%)
FIXTRUST-GS $N_r = 48$	PSNR ↑	14.96	16.49	**17.93**	**13.45**	**20.77**	**20.83**	**17.97**	17.85	**13.28**
	SSIM ↑	0.302	0.531	**0.545**	**0.255**	0.547	**0.692**	**0.627**	**0.302**	**0.345**
	LPIPS ↓	0.624	0.401	0.340	0.657	0.360	0.188	0.343	0.614	0.666
	#GS ↓	651 (↓42.6%)	460 (↓11.4%)	249 (↓9.3%)	516 (↓46.9%)	491 (↓32.7%)	343 (↓6.0%)	**222 (↓17.1%)**	**734 (↓37.8%)**	**356 (↓52.5%)**

Train 6 result. At $N_r = 48$, FIXTRUST-GS changes PSNR by +1.23 dB and SSIM by +0.053 over Train-only with 35.0% fewer Gaussians; LPIPS increases by 0.024. Macro PSNR for $N_r = 12/24/36/48$ is 16.62/16.93/17.04/17.06 dB. The scene-wise PSNR optimum spans $N_r = 36/48$, while mean Gaussian count changes from 499.2k to 446.8k.

Table 23: Complete Diffusion Harmonizer weight sweep at Train 9, averaged equally over Mip-NeRF 360. The Train-only row is the paired $N_r = 48$ baseline.

Supervision	$N_r = 12$	$N_r = 24$	$N_r = 36$	$N_r = 48$
TRAIN-ONLY		17.22 / 0.478 0.382 / 862.4		
Global 1.0	17.66 / 0.489 0.430 / 861.6	17.92 / 0.493 0.442 / 886.7	17.99 / 0.493 0.449 / 890.6	18.02 / 0.493 0.454 / 896.7
Global 0.9	17.70 / 0.491 0.428 / 827.8	17.93 / 0.494 0.441 / 839.6	17.98 / 0.494 0.449 / 836.6	18.04 / 0.494 0.452 / 844.2
Global 0.8	17.67 / 0.491 0.427 / 795.4	17.94 / 0.496 0.439 / 789.8	18.00 / 0.495 0.446 / 784.7	18.06 / 0.496 0.450 / 787.9
Global 0.7	17.70 / 0.493 0.424 / 765.0	17.94 / 0.497 0.437 / 743.2	18.02 / 0.497 0.445 / 735.9	18.06 / 0.497 0.449 / 736.9
Global 0.6	17.69 / 0.494 0.422 / 732.2	17.94 / 0.499 0.435 / 701.9	18.02 / 0.498 0.443 / 689.1	18.08 / 0.498 0.447 / 689.5
Global 0.5	17.75 / 0.496 0.418 / 704.2	18.00 / 0.501 0.430 / 661.6	18.07 / 0.501 0.439 / 648.8	18.08 / 0.500 0.445 / 644.7
Global 0.4	**17.77** / 0.498 0.416 / 676.4	18.00 / 0.503 0.428 / 626.8	18.07 / 0.503 0.436 / 612.5	18.11 / 0.502 0.441 / 608.4
Global 0.3	17.75 / 0.499 0.412 / 652.5	18.02 / 0.506 0.422 / 595.5	18.09 / 0.505 0.431 / 581.0	18.12 / 0.504 0.437 / 577.1
Global 0.2	17.70 / 0.499 0.410 / 633.1	**18.02** / 0.508 0.417 / 572.2	18.11 / 0.508 0.425 / 557.2	18.13 / 0.508 0.430 / 553.3
Global 0.1	17.69 / **0.499** 0.404 / **617.9**	17.96 / 0.508 0.410 / **555.3**	18.07 / 0.510 0.417 / **541.1**	18.11 / 0.510 0.422 / **538.0**
FIXTRUST-GS	17.69 / 0.497 **0.403** / 624.6	18.01 / **0.508** **0.408** / 562.9	**18.13** / **0.511** **0.415** / 546.8	**18.20** / **0.513** **0.418** / 542.9

Table 24: Scene-wise Diffusion Harmonizer results at Train 9. Repair-view count varies for FIXTRUST-GS; Train-only is the paired $N_r = 48$ run.

Supervision	Metric	Bicycle	Bonsai	Counter	Flowers	Garden	Kitchen	Room	Stump	Treehill
	PSNR ↑	15.18	17.62	17.93	12.81	20.94	20.69	19.13	16.77	13.89
	SSIM ↑	0.303	0.598	0.555	0.249	0.611	0.719	0.649	0.272	0.345
	LPIPS ↓	0.527	0.309	0.292	0.554	0.265	0.161	0.293	0.475	0.564
	#GS ↓	1485 (ref.)	630 (ref.)	299 (ref.)	1152 (ref.)	803 (ref.)	472 (ref.)	370 (ref.)	1445 (ref.)	1105 (ref.)
FIXTRUST-GS $N_r = 12$	PSNR ↑	15.90	17.75	18.26	13.43	21.46	21.18	19.41	17.72	14.07
	SSIM ↑	0.338	0.607	0.573	0.271	0.614	0.735	0.667	0.306	0.366
	LPIPS ↓	**0.556**	0.314	0.296	**0.606**	**0.295**	0.153	0.289	**0.531**	**0.585**
	#GS ↓	1176 (↓20.8%)	538 (↓14.6%)	269 (↓9.9%)	719 (↓37.6%)	528 (↓34.3%)	**417 (↓11.7%)**	316 (↓14.7%)	1009 (↓30.2%)	649 (↓41.3%)
FIXTRUST-GS $N_r = 24$	PSNR ↑	16.26	18.10	18.53	13.69	21.66	21.49	19.61	18.26	14.45
	SSIM ↑	0.351	0.625	0.588	0.278	**0.615**	0.744	0.677	0.321	0.378
	LPIPS ↓	0.567	0.307	**0.291**	0.619	0.307	0.148	0.288	0.552	0.591
	#GS ↓	914 (↓38.4%)	**537 (↓14.8%)**	264 (↓11.8%)	642 (↓44.3%)	516 (↓35.7%)	419 (↓11.2%)	300 (↓18.9%)	897 (↓37.9%)	576 (↓47.9%)
FIXTRUST-GS $N_r = 36$	PSNR ↑	16.33	18.19	18.79	**13.92**	21.70	21.54	19.83	18.25	14.67
	SSIM ↑	0.350	0.628	0.595	**0.282**	0.614	0.746	0.681	0.322	0.385
	LPIPS ↓	0.580	0.307	0.292	0.630	0.316	**0.148**	0.287	0.572	0.599
	#GS ↓	851 (↓42.7%)	537 (↓14.7%)	**262 (↓12.3%)**	622 (↓46.0%)	514 (↓36.0%)	421 (↓10.9%)	291 (↓21.5%)	870 (↓39.8%)	554 (↓49.9%)
FIXTRUST-GS $N_r = 48$	PSNR ↑	**16.35**	**18.34**	**18.86**	13.85	**21.71**	**21.62**	**19.83**	**18.32**	**14.88**
	SSIM ↑	**0.351**	**0.634**	**0.596**	0.281	0.612	**0.747**	**0.682**	**0.323**	**0.389**
	LPIPS ↓	0.588	**0.302**	0.291	0.630	0.323	0.149	**0.287**	0.584	0.606
	#GS ↓	**839 (↓43.5%)**	540 (↓14.3%)	263 (↓11.8%)	**621 (↓46.1%)**	**513 (↓36.1%)**	418 (↓11.4%)	**289 (↓22.0%)**	**854 (↓40.9%)**	**549 (↓50.3%)**

Train 9 result. At $N_r = 48$, FIXTRUST-GS changes PSNR by +0.98 dB and SSIM by +0.035 over Train-only with 37.0% fewer Gaussians; LPIPS increases by 0.036. Macro PSNR for $N_r = 12/24/36/48$ is 17.69/18.01/18.13/18.20 dB. The scene-wise PSNR optimum spans $N_r = 36/48$, while mean Gaussian count changes from 624.6k to 542.9k.

C.3.4 Diffusion Harmonizer on Deep Blending

The complete matrix covers Deep Blending with $N_t \in \{3, 6, 9\}$, $N_r \in \{12, 24, 36, 48\}$, and $\lambda_g \in \{0.1, 0.2, \ldots, 1.0\}$. It follows the shared protocol in Appendix C.1, including nested repair-camera subsets and matched continuation branches within every evaluation cell.

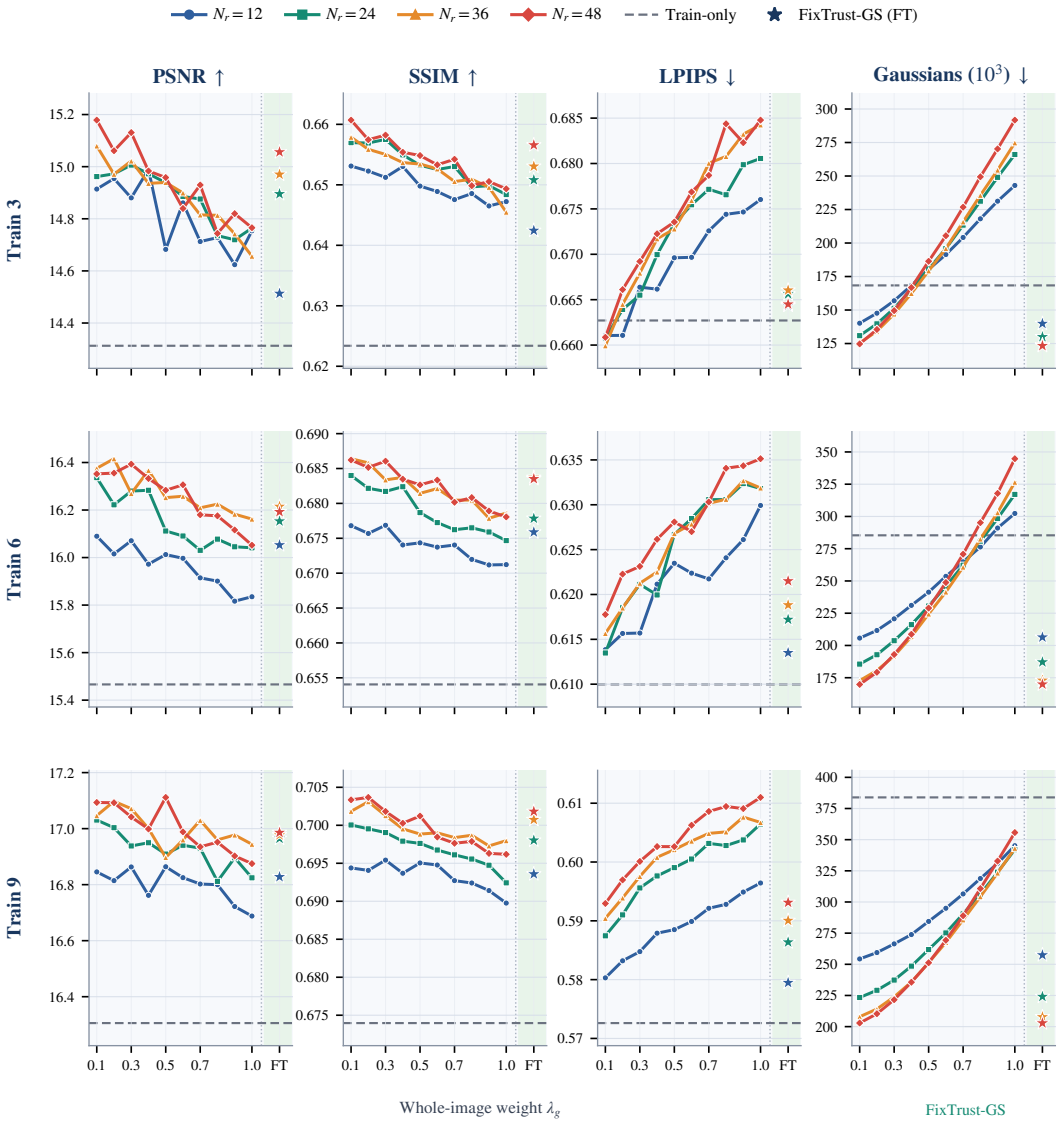

Figure 9: Complete Diffusion Harmonizer response on Deep Blending. Rows fix the captured training-view budget and colored curves fix N_r. Curves sweep the global weight $\lambda_g = 0.1$–1.0; stars denote FixTrust-GS, and the gray dashed line is Train-only. Values are scene-macro averages at step 8,999.

Result summary. At $N_r = 48$, FixTrust-GS reaches 15.06/16.19/16.99 dB for Train 3/6/9 while using 123.3/169.9/202.8 thousand Gaussians. Some uniform weights give higher PSNR, while FixTrust-GS reduces LPIPS and Gaussian count relative to Global 1.0 at each training-view budget. Smaller global weights yield different quality and size trade-offs. The response curves and scene-wise tables below report these variations.

Table 25: Complete Diffusion Harmonizer weight sweep at Train 3, averaged equally over Deep Blending. The Train-only row is the paired $N_r = 48$ baseline.

Supervision	$N_r = 12$	$N_r = 24$	$N_r = 36$	$N_r = 48$
TRAIN-ONLY		14.31 / 0.623 0.663 / 168.4		
Global 1.0	14.75 / 0.647 0.676 / 243.0	14.76 / 0.648 0.681 / 266.1	14.66 / 0.645 0.684 / 274.8	14.77 / 0.649 0.685 / 291.7
Global 0.9	14.62 / 0.647 0.675 / 231.2	14.72 / 0.650 0.680 / 249.0	14.74 / 0.650 0.683 / 254.4	14.82 / 0.651 0.682 / 270.2
Global 0.8	14.73 / 0.649 0.674 / 218.1	14.74 / 0.650 0.677 / 231.2	14.81 / 0.651 0.681 / 235.6	14.74 / 0.650 0.684 / 249.4
Global 0.7	14.71 / 0.648 0.673 / 204.1	14.88 / 0.653 0.677 / 213.3	14.81 / 0.651 0.680 / 215.8	14.93 / 0.654 0.679 / 226.8
Global 0.6	14.86 / 0.649 0.670 / 191.4	14.89 / 0.653 0.675 / 195.9	14.90 / 0.653 0.676 / 195.9	14.84 / 0.653 0.677 / 205.4
Global 0.5	14.68 / 0.650 0.670 / 180.3	14.94 / 0.653 0.673 / 180.7	14.94 / 0.653 0.673 / 179.2	14.96 / 0.655 0.674 / 186.4
Global 0.4	**14.99** / 0.653 0.666 / 168.4	14.97 / 0.655 0.670 / 165.4	14.94 / 0.654 0.672 / 162.4	14.98 / 0.655 0.672 / 166.8
Global 0.3	14.88 / 0.651 0.666 / 156.9	**15.01** / **0.658** 0.665 / 151.3	15.02 / 0.655 0.668 / 147.1	15.13 / 0.658 0.669 / 149.5
Global 0.2	14.95 / 0.652 0.661 / 147.6	14.97 / 0.657 0.664 / 139.7	14.97 / 0.656 0.664 / 134.3	15.06 / 0.657 0.666 / 135.5
Global 0.1	14.91 / **0.653** **0.661** / 140.2	14.96 / 0.657 **0.661** / 130.9	**15.08** / **0.658** **0.660** / 125.3	**15.18** / **0.661** **0.661** / 124.7
FIXTRUST-GS	14.51 / 0.642 0.664 / **139.9**	14.90 / 0.651 0.666 / **130.0**	14.97 / 0.653 0.666 / **123.8**	15.06 / 0.657 0.664 / **123.3**

Table 26: Scene-wise Diffusion Harmonizer results at Train 3. Repair-view count varies for FIXTRUST-GS; Train-only is the paired $N_r = 48$ run.

Supervision	Metric	DrJohnson	Playroom	Bedroom	Boats	Bridge	CreepyAttic
	PSNR ↑	16.19	13.53	14.52	10.37	14.94	16.33
	SSIM ↑	0.651	0.643	0.628	0.664	0.529	0.625
	LPIPS ↓	0.675	0.645	0.615	0.600	0.667	0.775
	#GS ↓	155 (ref.)	125 (ref.)	124 (ref.)	236 (ref.)	126 (ref.)	243 (ref.)
FIXTRUST-GS $N_r = 12$	PSNR ↑	17.10	14.68	14.48	11.20	14.34	15.29
	SSIM ↑	0.681	0.684	0.634	0.704	0.542	0.609
	LPIPS ↓	0.666	0.638	0.606	0.589	**0.693**	**0.794**
	#GS ↓	122 (↓21.2%)	112 (↓10.9%)	**126** (↑1.8%)	179 (↓24.4%)	98 (↓22.6%)	203 (↓16.6%)
FIXTRUST-GS $N_r = 24$	PSNR ↑	17.02	14.85	15.03	**11.70**	15.13	**15.63**
	SSIM ↑	0.681	0.686	0.642	**0.721**	0.561	0.614
	LPIPS ↓	0.669	0.637	0.613	**0.580**	0.693	0.801
	#GS ↓	122 (↓21.3%)	106 (↓15.3%)	101 (↓18.4%)	179 (↓24.2%)	90 (↓29.0%)	182 (↓25.4%)
FIXTRUST-GS $N_r = 36$	PSNR ↑	**17.17**	15.10	15.16	11.52	15.46	15.42
	SSIM ↑	0.684	0.692	0.644	0.718	0.567	**0.614**
	LPIPS ↓	0.666	0.633	0.611	0.584	0.696	0.806
	#GS ↓	**122** (↓21.6%)	**99** (↓20.7%)	97 (↓21.3%)	**167** (↓29.4%)	88 (↓30.0%)	169 (↓30.5%)
FIXTRUST-GS $N_r = 48$	PSNR ↑	17.13	**15.58**	**15.22**	11.56	**15.48**	15.36
	SSIM ↑	**0.686**	**0.704**	**0.649**	0.721	**0.567**	0.614
	LPIPS ↓	**0.665**	**0.630**	**0.603**	0.583	0.697	0.809
	#GS ↓	122 (↓21.5%)	101 (↓19.8%)	**95** (↓23.1%)	169 (↓28.7%)	**87** (↓31.4%)	**167** (↓31.2%)

Train 3 result. At $N_r = 48$, FIXTRUST-GS changes PSNR by +0.74 dB and SSIM by +0.033 over Train-only with 26.8% fewer Gaussians; LPIPS increases by 0.002. Macro PSNR for $N_r = 12/24/36/48$ is 14.51/14.90/14.97/15.06 dB. The scene-wise PSNR optimum spans $N_r = 24/36/48$, while mean Gaussian count changes from 139.9k to 123.3k.

Table 27: Complete Diffusion Harmonizer weight sweep at Train 6, averaged equally over Deep Blending. The Train-only row is the paired $N_r = 48$ baseline.

Supervision	$N_r = 12$	$N_r = 24$	$N_r = 36$	$N_r = 48$
TRAIN-ONLY			15.47 / 0.654 0.610 / 285.4	
Global 1.0	15.84 / 0.671 0.630 / 302.3	16.04 / 0.675 0.632 / 317.1	16.16 / 0.679 0.632 / 326.5	16.05 / 0.678 0.635 / 344.7
Global 0.9	15.82 / 0.671 0.626 / 291.0	16.05 / 0.676 0.632 / 298.4	16.18 / 0.678 0.633 / 302.9	16.12 / 0.679 0.634 / 317.9
Global 0.8	15.90 / 0.672 0.624 / 276.4	16.08 / 0.677 0.631 / 280.8	16.22 / 0.680 0.631 / 282.5	16.18 / 0.681 0.634 / 295.2
Global 0.7	15.91 / 0.674 0.622 / 264.8	16.03 / 0.676 0.631 / 262.2	16.21 / 0.680 0.630 / 260.7	16.18 / 0.680 0.630 / 270.8
Global 0.6	16.00 / 0.674 0.622 / 253.6	16.09 / 0.677 0.628 / 245.9	16.26 / 0.682 0.628 / 241.5	16.31 / 0.683 0.627 / 248.8
Global 0.5	16.01 / 0.674 0.623 / 241.3	16.11 / 0.679 0.627 / 230.5	16.25 / 0.681 0.627 / 224.4	16.28 / 0.683 0.628 / 229.0
Global 0.4	15.97 / 0.674 0.621 / 231.1	16.28 / 0.682 0.620 / 216.1	16.37 / 0.684 0.623 / 206.8	16.33 / 0.683 0.626 / 208.7
Global 0.3	16.07 / **0.677** 0.616 / 220.6	16.28 / 0.682 0.621 / 203.7	16.27 / 0.683 0.621 / 191.9	**16.39** / 0.686 0.623 / 193.0
Global 0.2	16.02 / 0.676 0.616 / 211.6	16.22 / 0.682 0.619 / 192.8	**16.42** / 0.686 0.618 / 180.6	16.36 / 0.685 0.622 / 179.2
Global 0.1	**16.09** / 0.677 0.614 / **205.7**	**16.34** / **0.684** **0.613** / **185.6**	16.38 / **0.686** **0.616** / **172.8**	16.35 / **0.686** **0.618** / **169.8**
FIXTRUST-GS	16.05 / 0.676 **0.614** / 206.4	16.15 / 0.678 0.617 / 187.1	16.21 / 0.684 0.619 / 172.9	16.19 / 0.684 0.622 / 169.9

Table 28: Scene-wise Diffusion Harmonizer results at Train 6. Repair-view count varies for FIXTRUST-GS; Train-only is the paired $N_r = 48$ run.

Supervision	Metric	DrJohnson	Playroom	Bedroom	Boats	Bridge	CreepyAttic
	PSNR ↑	16.14	15.72	14.79	14.68	15.07	16.41
	SSIM ↑	0.656	0.689	0.644	0.769	0.539	0.627
	LPIPS ↓	0.662	0.550	0.604	0.460	0.669	0.714
	#GS ↓	259 (ref.)	227 (ref.)	209 (ref.)	425 (ref.)	172 (ref.)	421 (ref.)
FIXTRUST-GS $N_r = 12$	PSNR ↑	16.94	16.90	15.31	14.98	**15.59**	16.59
	SSIM ↑	0.682	0.722	0.660	0.784	0.568	0.639
	LPIPS ↓	**0.656**	0.550	0.594	**0.472**	**0.687**	**0.723**
	#GS ↓	175 (↓32.6%)	176 (↓22.4%)	156 (↓25.4%)	337 (↓20.7%)	108 (↓37.3%)	288 (↓31.7%)
FIXTRUST-GS $N_r = 24$	PSNR ↑	17.08	**17.25**	15.57	15.11	15.46	16.45
	SSIM ↑	0.687	0.729	0.669	0.787	0.569	0.626
	LPIPS ↓	0.657	**0.545**	0.586	0.478	0.701	0.738
	#GS ↓	167 (↓35.6%)	161 (↓28.7%)	133 (↓36.3%)	319 (↓24.8%)	94 (↓45.0%)	248 (↓41.2%)
FIXTRUST-GS $N_r = 36$	PSNR ↑	**17.17**	17.09	15.82	15.12	15.21	**16.87**
	SSIM ↑	**0.691**	0.732	0.676	**0.790**	0.569	**0.644**
	LPIPS ↓	0.656	0.554	**0.582**	0.480	0.710	0.731
	#GS ↓	163 (↓37.1%)	151 (↓33.3%)	130 (↓37.7%)	273 (↓35.8%)	93 (↓46.1%)	228 (↓45.9%)
FIXTRUST-GS $N_r = 48$	PSNR ↑	17.03	17.19	**15.89**	**15.16**	15.22	16.67
	SSIM ↑	0.689	**0.733**	**0.680**	0.790	**0.571**	0.637
	LPIPS ↓	0.661	0.551	0.582	0.485	0.713	0.736
	#GS ↓	**163** (↓37.3%)	**149** (↓34.1%)	**127** (↓39.2%)	**263** (↓38.0%)	**92** (↓46.6%)	**226** (↓46.4%)

Train 6 result. At $N_r = 48$, FIXTRUST-GS changes PSNR by +0.73 dB and SSIM by +0.029 over Train-only with 40.5% fewer Gaussians; LPIPS increases by 0.012. Macro PSNR for $N_r = 12/24/36/48$ is 16.05/16.15/16.21/16.19 dB. The scene-wise PSNR optimum spans $N_r = 12/24/36/48$, while mean Gaussian count changes from 206.4k to 169.9k.

Table 29: Complete Diffusion Harmonizer weight sweep at Train 9, averaged equally over Deep Blending. The Train-only row is the paired $N_r = 48$ baseline.

Supervision	$N_r = 12$	$N_r = 24$	$N_r = 36$	$N_r = 48$
TRAIN-ONLY		16.31 / 0.674 0.573 / 383.9		
Global 1.0	16.69 / 0.690 0.596 / 345.2	16.82 / 0.692 0.606 / 342.0	16.94 / 0.698 0.607 / 343.1	16.88 / 0.696 0.611 / 355.7
Global 0.9	16.72 / 0.691 0.595 / 331.0	16.90 / 0.695 0.604 / 323.9	16.98 / 0.697 0.608 / 323.2	16.90 / 0.696 0.609 / 332.8
Global 0.8	16.80 / 0.692 0.593 / 318.8	16.81 / 0.696 0.603 / 306.2	16.96 / 0.699 0.605 / 304.2	16.95 / 0.698 0.609 / 310.8
Global 0.7	16.80 / 0.693 0.592 / 306.5	16.93 / 0.696 0.603 / 290.5	17.03 / 0.698 0.605 / 285.7	16.93 / 0.698 0.609 / 289.0
Global 0.6	16.82 / 0.695 0.590 / 295.0	16.94 / 0.697 0.601 / 275.2	16.96 / 0.699 0.604 / 267.5	16.99 / 0.698 0.606 / 269.3
Global 0.5	**16.86** / 0.695 0.588 / 284.4	16.91 / 0.698 0.599 / 261.8	16.90 / 0.699 0.602 / 251.3	**17.11** / 0.701 0.603 / 251.2
Global 0.4	16.76 / 0.694 0.588 / 273.9	16.95 / 0.698 0.598 / 248.5	17.00 / 0.700 0.601 / 235.8	17.00 / 0.700 0.603 / 235.6
Global 0.3	16.86 / **0.695** 0.585 / 266.4	16.94 / 0.699 0.596 / 237.4	17.07 / 0.701 0.598 / 223.9	17.04 / 0.702 0.600 / 221.6
Global 0.2	16.81 / 0.694 0.583 / 259.4	17.00 / 0.700 0.591 / 229.2	**17.10** / **0.703** 0.594 / 214.1	17.09 / **0.704** 0.597 / 210.3
Global 0.1	16.85 / 0.694 0.580 / **254.3**	**17.03** / **0.700** 0.587 / **223.4**	17.05 / 0.702 0.590 / 207.9	17.09 / 0.703 **0.593** / 202.9
FIXTRUST-GS	16.83 / 0.694 **0.579** / 257.4	16.96 / 0.698 **0.586** / 224.1	16.98 / 0.701 **0.590** / **207.7**	16.99 / 0.702 0.593 / **202.8**

Table 30: Scene-wise Diffusion Harmonizer results at Train 9. Repair-view count varies for FIXTRUST-GS; Train-only is the paired $N_r = 48$ run.

Supervision	Metric	DrJohnson	Playroom	Bedroom	Boats	Bridge	CreepyAttic
	PSNR ↑	16.39	16.03	15.62	16.41	15.84	17.54
	SSIM ↑	0.658	0.698	0.673	0.794	0.577	0.644
	LPIPS ↓	0.648	0.530	0.567	0.421	0.634	0.636
	#GS ↓	392 (ref.)	314 (ref.)	274 (ref.)	513 (ref.)	172 (ref.)	638 (ref.)
FIXTRUST-GS $N_r = 12$	PSNR ↑	17.40	17.23	16.02	16.50	**16.22**	17.61
	SSIM ↑	0.687	0.729	0.690	0.803	0.596	0.657
	LPIPS ↓	**0.644**	**0.531**	0.553	**0.444**	**0.658**	**0.647**
	#GS ↓	224 (↓42.7%)	222 (↓29.3%)	201 (↓26.7%)	396 (↓22.8%)	112 (↓34.6%)	389 (↓39.1%)
FIXTRUST-GS $N_r = 24$	PSNR ↑	17.53	17.41	16.19	**16.60**	16.17	17.88
	SSIM ↑	0.691	0.734	0.694	**0.806**	0.599	0.664
	LPIPS ↓	0.651	0.537	**0.547**	0.454	0.671	0.658
	#GS ↓	198 (↓49.5%)	200 (↓36.4%)	177 (↓35.4%)	363 (↓29.3%)	100 (↓42.1%)	308 (↓51.8%)
FIXTRUST-GS $N_r = 36$	PSNR ↑	17.64	17.63	**16.26**	16.51	15.88	17.94
	SSIM ↑	0.695	0.742	**0.697**	0.805	0.600	0.666
	LPIPS ↓	0.656	0.537	0.549	0.461	0.676	0.661
	#GS ↓	190 (↓51.6%)	190 (↓39.6%)	173 (↓36.9%)	323 (↓37.1%)	98 (↓43.2%)	274 (↓57.1%)
FIXTRUST-GS $N_r = 48$	PSNR ↑	**17.65**	**17.67**	16.21	16.46	15.96	**17.97**
	SSIM ↑	**0.697**	**0.743**	0.696	0.805	**0.602**	**0.668**
	LPIPS ↓	0.656	0.538	0.554	0.468	0.680	0.662
	#GS ↓	**185** (↓52.7%)	**188** (↓40.2%)	**167** (↓39.1%)	**313** (↓39.0%)	**97** (↓43.6%)	**267** (↓58.2%)

Train 9 result. At $N_r = 48$, FIXTRUST-GS changes PSNR by +0.68 dB and SSIM by +0.028 over Train-only with 47.2% fewer Gaussians; LPIPS increases by 0.020. Macro PSNR for $N_r = 12/24/36/48$ is 16.83/16.96/16.98/16.99 dB. The scene-wise PSNR optimum spans $N_r = 12/24/36/48$, while mean Gaussian count changes from 257.4k to 202.8k.

C.4 Representative Qualitative Comparisons

These cases extend Section 4.2 across repair-model–dataset combinations and Train 3/6/9. Rows fix scene and camera; columns show the initial render, repair, weight overlay, final reconstructions, and ground truth. Labels give per-view PSNR; warmer overlays indicate higher weights.

C.4.1 Difix3D on Mip-NeRF 360: Train 3



Figure 10: Aligned qualitative comparisons for Difix3D on Mip-NeRF 360 with Train 3. Columns show the initial render, repair output, weight overlay, Train-only, global weighting, FixTrust-GS, and ground truth.

C.4.2 DIFIX3D ON MIP-NERF 360: TRAIN 6

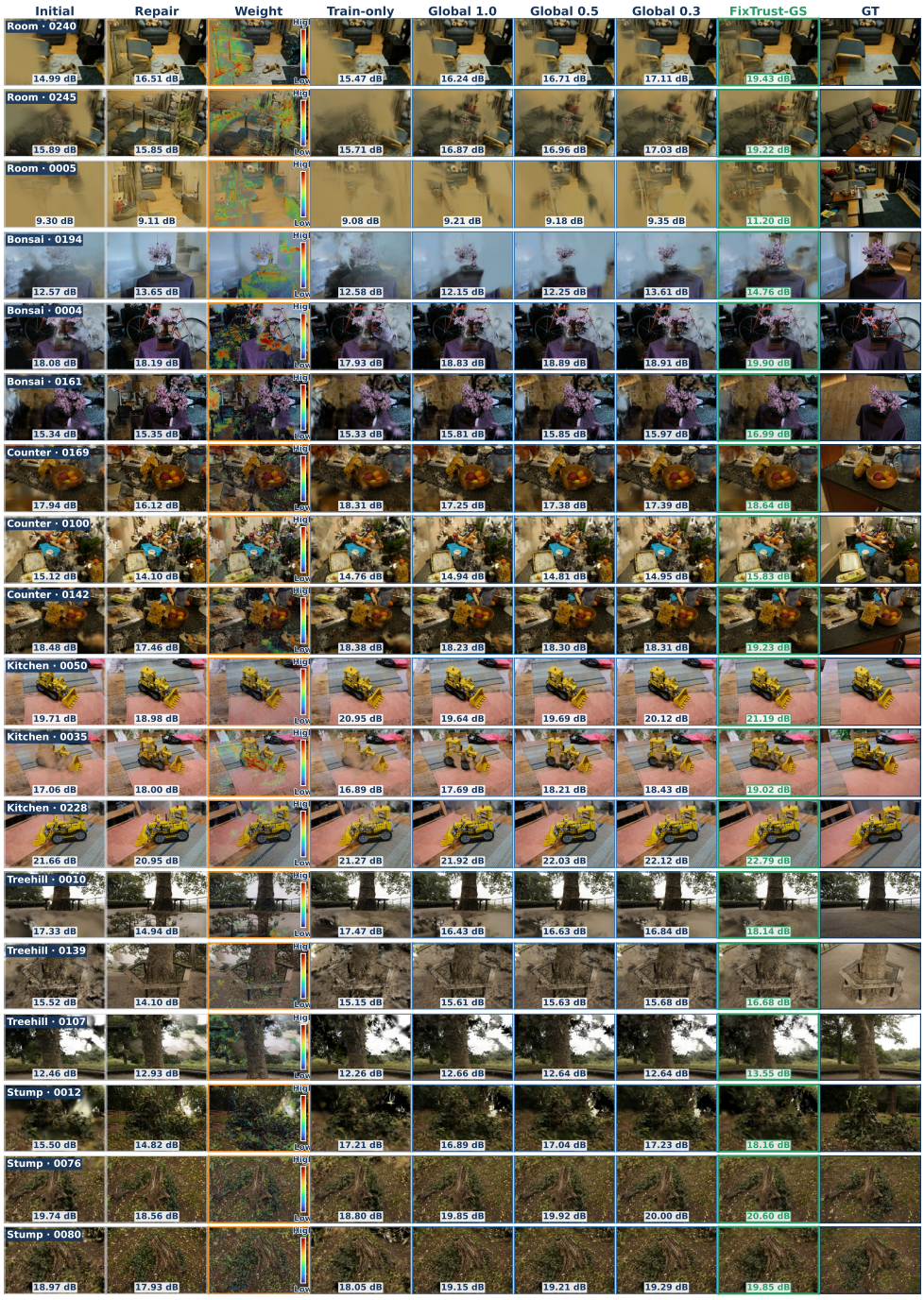

Figure 11: Aligned qualitative comparisons for Difix3D on Mip-NeRF 360 with Train 6. Columns show the initial render, repair output, weight overlay, Train-only, global weighting, FIXTRUST-GS, and ground truth.

C.4.3 DIFIX3D ON MIP-NERF 360: TRAIN 9



Figure 12: Aligned qualitative comparisons for Difix3D on Mip-NeRF 360 with Train 9. Columns show the initial render, repair output, weight overlay, Train-only, global weighting, FIXTRUST-GS, and ground truth.

C.4.4 DIFIX3D ON DEEP BLENDING: TRAIN 3

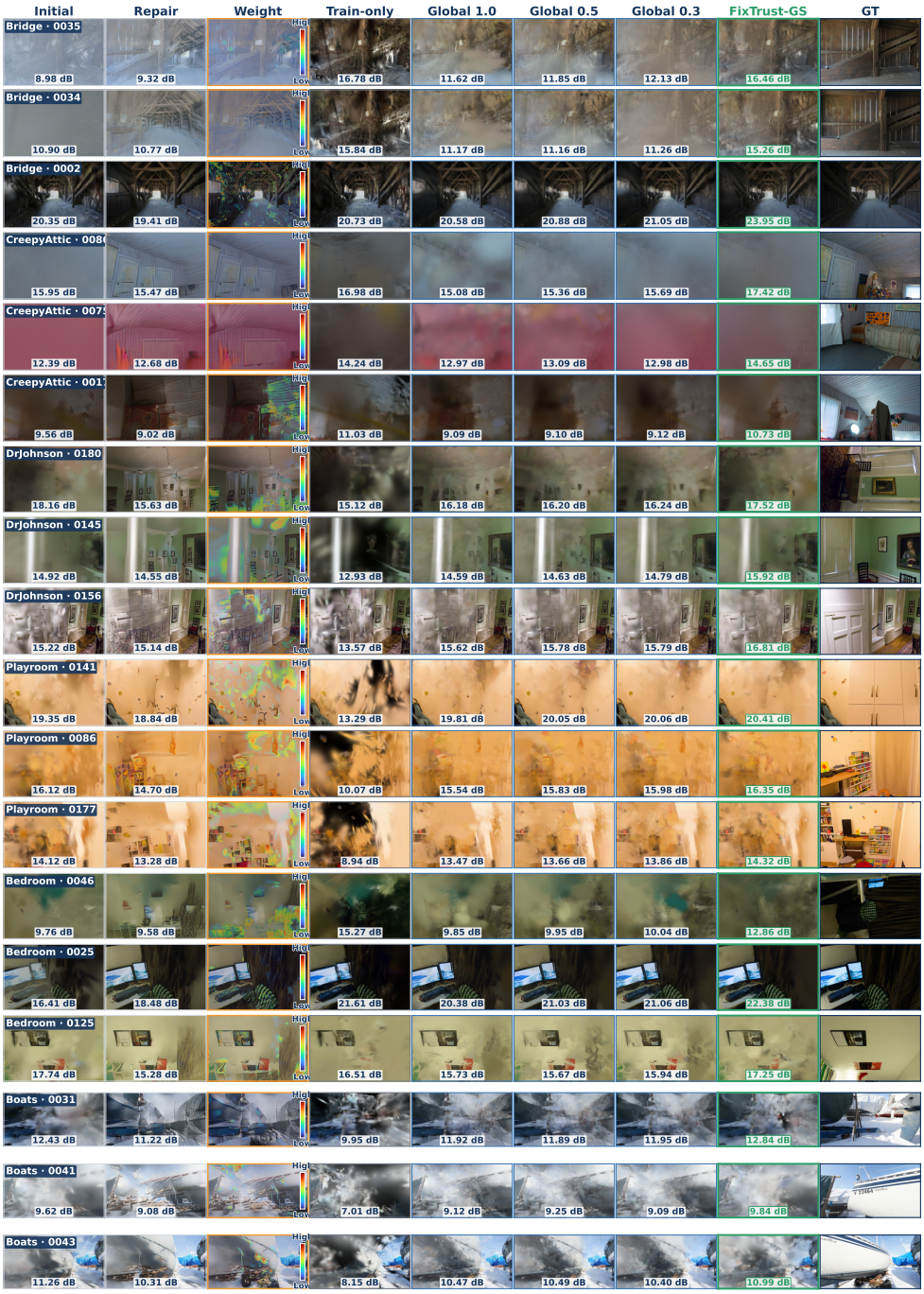

Figure 13: Aligned qualitative comparisons for Difix3D on Deep Blending with Train 3. Columns show the initial render, repair output, weight overlay, Train-only, global weighting, FIXTRUST-GS, and ground truth.

C.4.5 DIFIX3D ON DEEP BLENDING: TRAIN 6

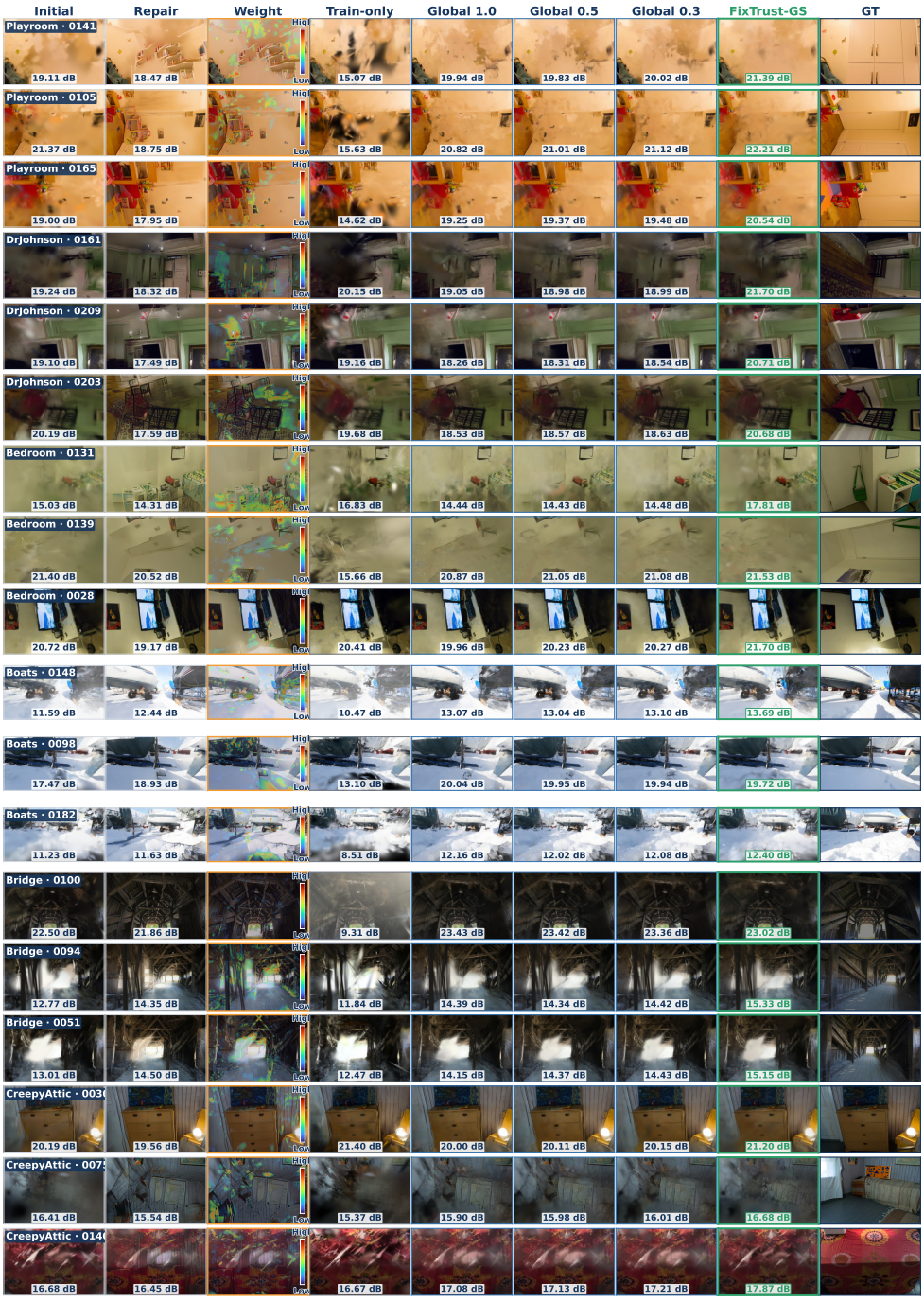

Figure 14: Aligned qualitative comparisons for Difix3D on Deep Blending with Train 6. Columns show the initial render, repair output, weight overlay, Train-only, global weighting, FIXTRUST-GS, and ground truth.

C.4.6 DIFIX3D ON DEEP BLENDING: TRAIN 9

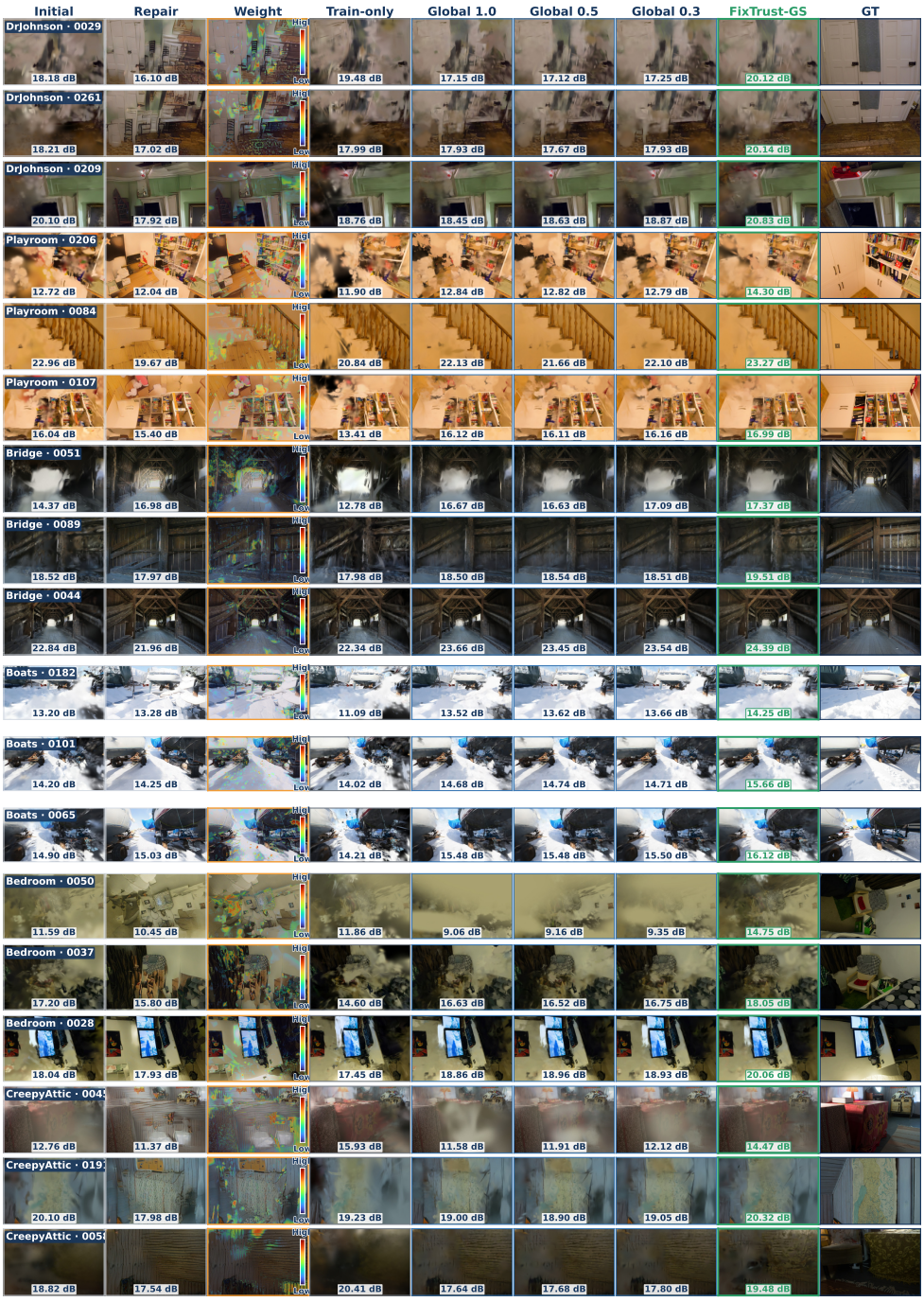

Figure 15: Aligned qualitative comparisons for Difix3D on Deep Blending with Train 9. Columns show the initial render, repair output, weight overlay, Train-only, global weighting, FIXTRUST-GS, and ground truth.

C.4.7 Diffusion Harmonizer on Mip-NeRF 360: Train 3



Figure 16: Aligned qualitative comparisons for Diffusion Harmonizer on Mip-NeRF 360 with Train 3. Columns show the initial render, repair output, weight overlay, Train-only, global weighting, FixTrust-GS, and ground truth.

C.4.8 Diffusion Harmonizer on Mip-NeRF 360: Train 6

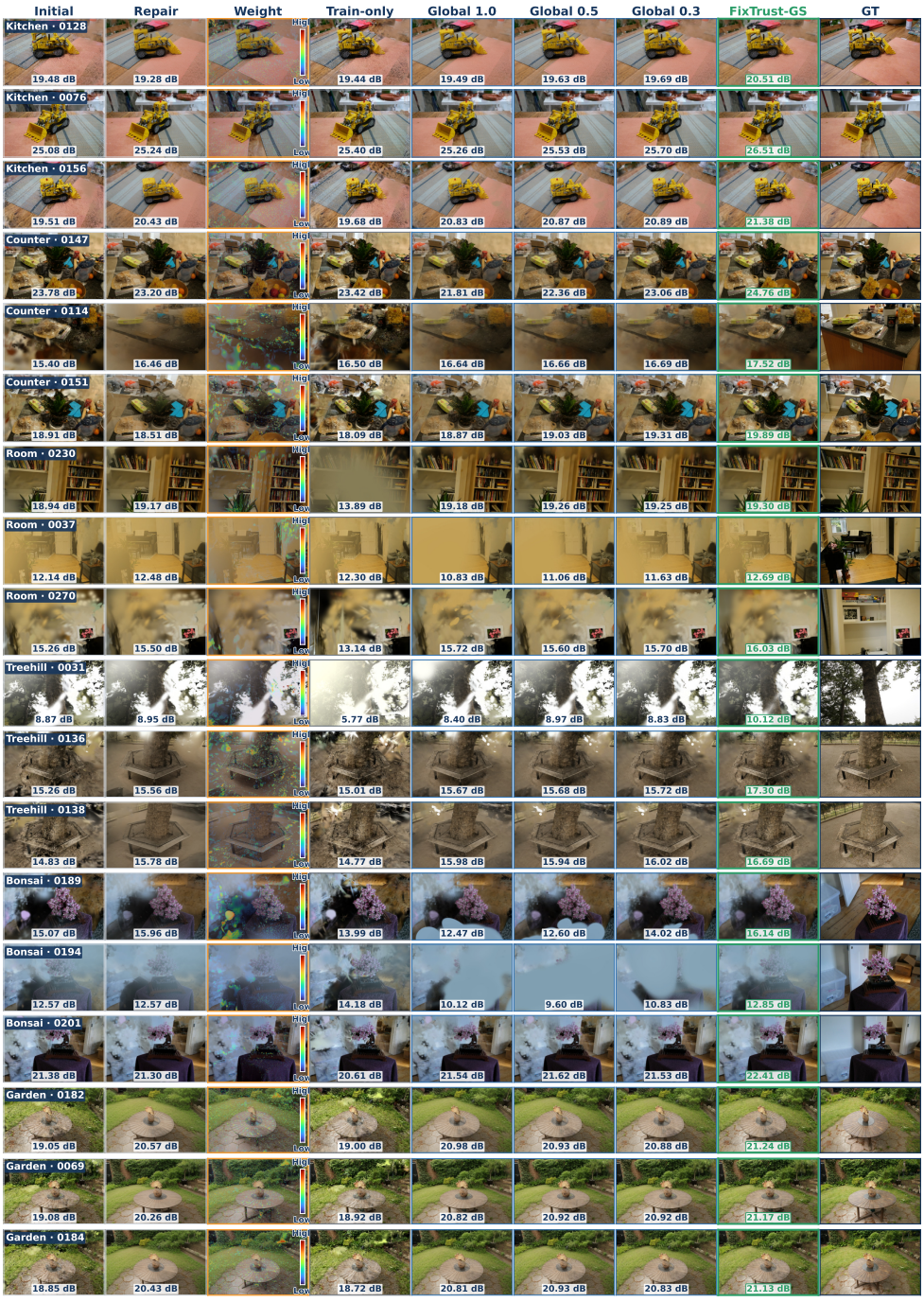

Figure 17: Aligned qualitative comparisons for Diffusion Harmonizer on Mip-NeRF 360 with Train 6. Columns show the initial render, repair output, weight overlay, Train-only, global weighting, FixTrust-GS, and ground truth.

C.4.9 Diffusion Harmonizer on Mip-NeRF 360: Train 9

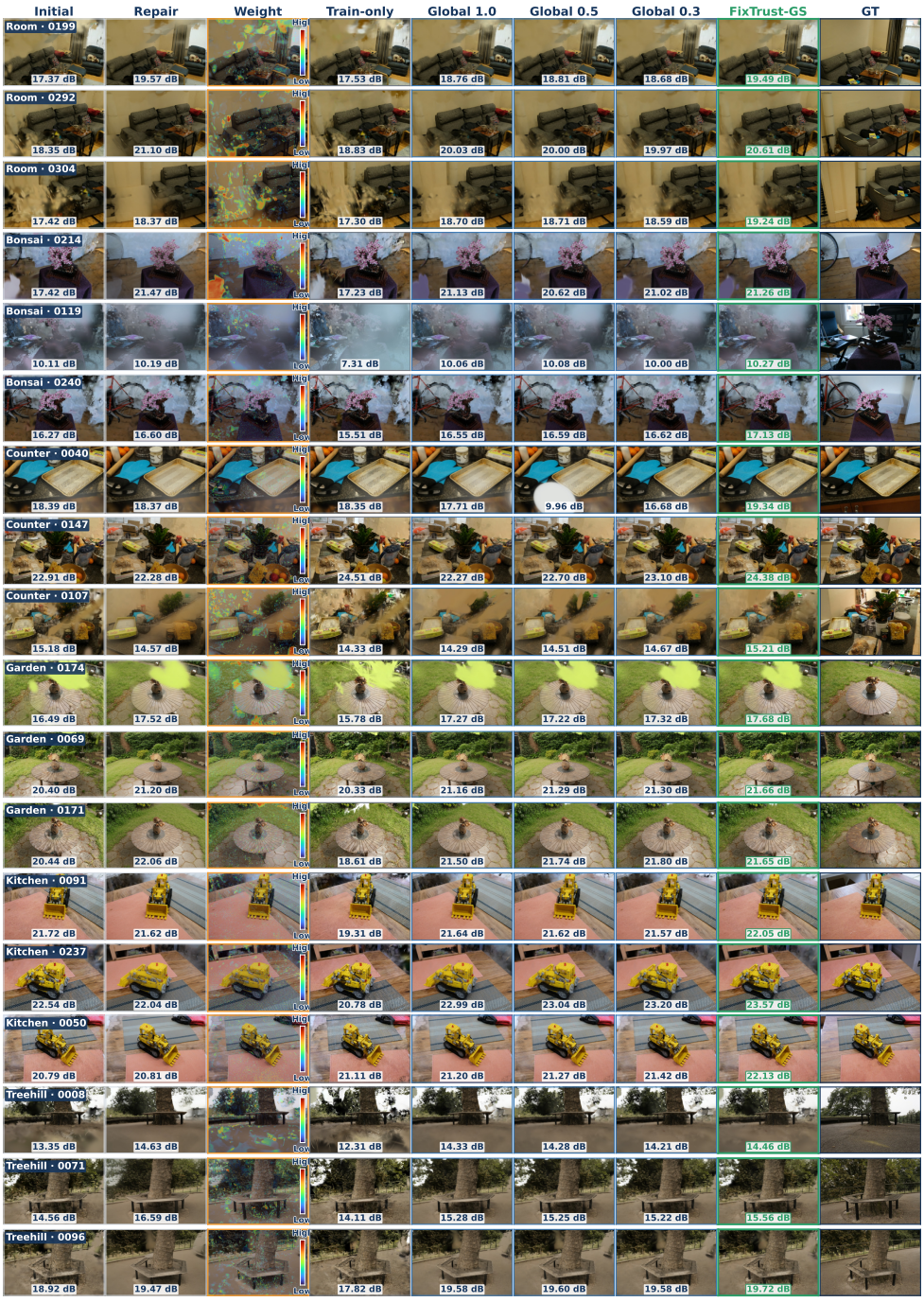

Figure 18: Aligned qualitative comparisons for Diffusion Harmonizer on Mip-NeRF 360 with Train 9. Columns show the initial render, repair output, weight overlay, Train-only, global weighting, FixTrust-GS, and ground truth.

C.4.10 Diffusion Harmonizer on Deep Blending: Train 3



Figure 19: Aligned qualitative comparisons for Diffusion Harmonizer on Deep Blending with Train 3. Columns show the initial render, repair output, weight overlay, Train-only, global weighting, FixTrust-GS, and ground truth.

C.4.11 Diffusion Harmonizer on Deep Blending: Train 6



Figure 20: Aligned qualitative comparisons for Diffusion Harmonizer on Deep Blending with Train 6. Columns show the initial render, repair output, weight overlay, Train-only, global weighting, FixTrust-GS, and ground truth.

C.4.12 Diffusion Harmonizer on Deep Blending: Train 9

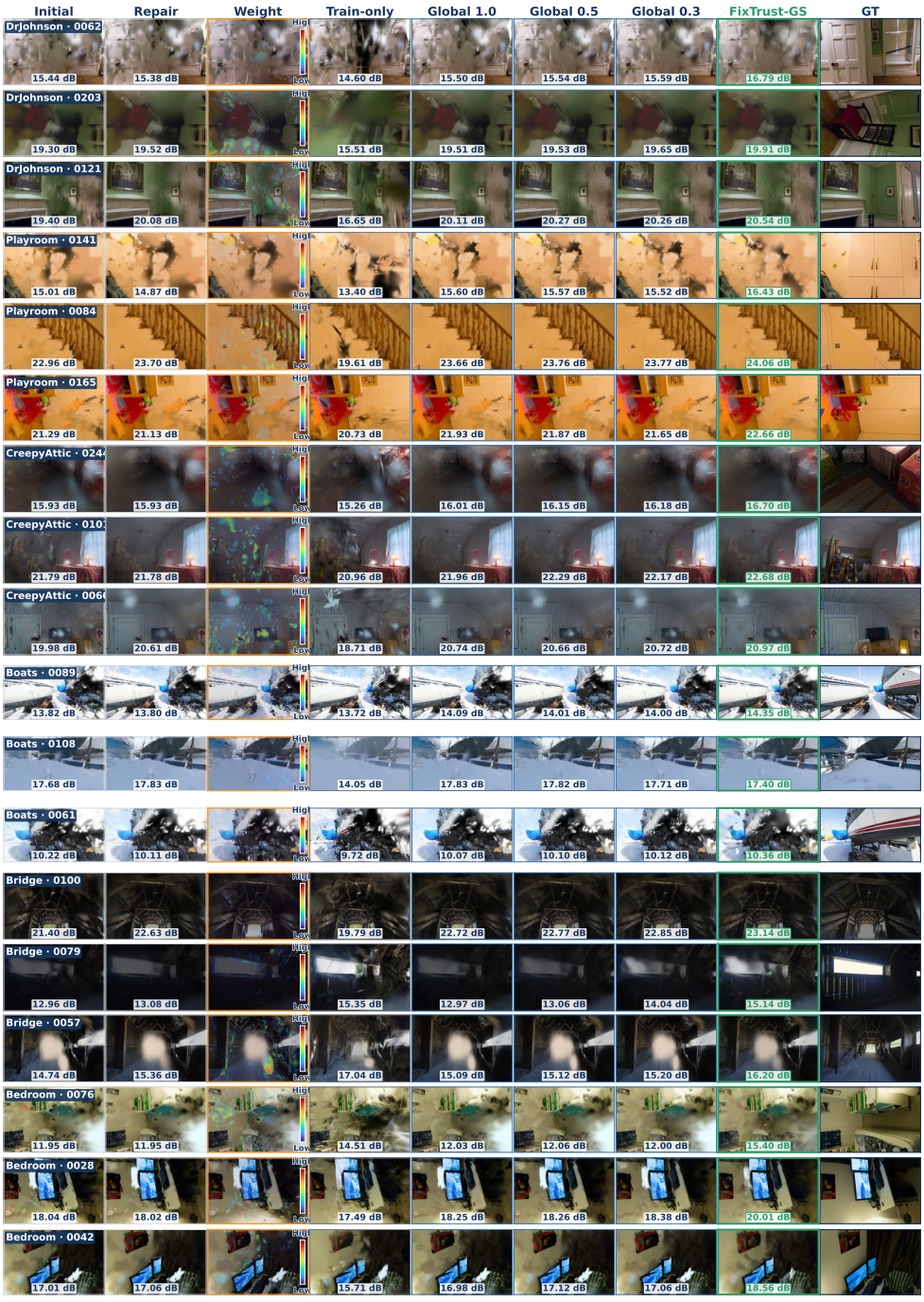

Figure 21: Aligned qualitative comparisons for Diffusion Harmonizer on Deep Blending with Train 9. Columns show the initial render, repair output, weight overlay, Train-only, global weighting, FixTrust-GS, and ground truth.