TOWARDS GENERAL REAL-WORLD IMAGE RESTORATION: AIGC-SUPERVISED LEARNING WITH KNOWLEDGE DISTILLATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Real-world image restoration must accommodate complex variations across scenes while supporting deployment under limited computational resources. However, acquiring paired data is costly and challenging, and synthetic degradations cannot fully capture real image formation processes, limiting the generalization of restoration models to real-world scenarios. We propose a real-world image restoration framework that integrates generative supervision, a perturbation-aware structural correction module (PSR), and knowledge distillation without requiring real clean targets. An input-anchored local risk-constrained selection strategy combines local content constraints, candidate disagreement, and high-risk region aggregation to select a pseudo-GT target with lower proxy content risk from candidates of similar perceptual quality. The selected target is paired with the degraded input to train a restoration teacher. To address variations in pseudo-GT targets across generative models or generation batches, we introduce PSR after the restoration teacher. Using selected alternative targets to guide local teacher probing, PSR learns a shared correction mapping through restoration supervision, response-weighted consistency, and input content constraints. The composite teacher incorporating PSR is then distilled into a student model. We also construct the XSMax dataset with scene-disjoint training, validation, and test splits. Extensive experiments demonstrate that our method improves the perceptual quality and content fidelity of restored images while enhancing output stability under variations in generative supervision. The proposed framework provides a practical approach to cross-scene image restoration without real clean targets and deployment in resource-constrained environments.

1 INTRODUCTION

Real-world image restoration aims to recover clear images from observations affected by complex degradations while faithfully preserving scene content. For a fixed imaging device, the sensor and imaging pipeline remain relatively stable, but variations in scenes, illumination, and capture conditions can shift the input distribution. Limited on-device computational and memory resources further constrain model size. We therefore study cross-scene real-world image restoration for a single device, balancing content fidelity, perceptual quality, and model compactness for on-device deployment.

Collecting precisely aligned real-world degraded–clean image pairs is costly (Flepp et al., 2024; Kong et al., 2026), while synthetic degradations may fail to capture actual imaging processes (Zhang et al., 2021; Wang et al., 2021). Generative multimodal foundation models offer an alternative: GGT-100K (Kong et al., 2026) pairs generated restoration targets with real degraded inputs. However, generated targets may alter text, boundaries, and local textures, potentially leading the restoration model to learn mappings inconsistent with the input content. To address this issue, we combine input content constraints, local disagreement among candidates, and aggregation over high-risk regions to select a single pseudo-ground-truth (pseudo-GT) target with a lower proxy score for content risk among candidates of comparable perceptual quality. The selected target is paired with its degraded input to train a restoration teacher.

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.

Selection alone cannot ensure consistent pseudo-GT targets across different generative models or generation batches from the same model. Residual variations in supervision may propagate through teacher training and affect output quality and details. To mitigate this effect, we introduce a Perturbation-aware Structural Refinement (PSR) module after the frozen restoration teacher. Alternative targets obtained through four independent rounds of generation and selection guide local probing of the teacher. PSR measures pixel-level responses and changes in inter-region relationships, and learns a shared correction mapping through restoration supervision, response-weighted consistency, and input content constraints.

Building on prior work on efficient restoration and teacher-guided learning (Chen et al., 2022; 2024b; Zhou et al., 2025), we freeze the composite teacher comprising the restoration teacher and PSR, and distill it into a compact student model.

Our main contributions are as follows:

- **A dataset and evaluation protocol for cross-scene real-world image restoration.** We construct XSMax, comprising 9,900 real-world images at a resolution of 1024×1024, spanning 8 coarse scene categories and 66 fine-grained scene categories. The dataset includes mixed degradations arising under real capture conditions, including noise, blur, underexposure, and artifacts. We adopt scene-disjoint splits to evaluate restoration on unseen scenes without real clean supervision.
- **Local risk constraints for generative supervision.** We combine input-anchored content constraints, local disagreement among candidates, and aggregation over high-risk regions to incorporate local content risk into the construction of restoration supervision.
- **Perturbation-aware structural refinement under supervision variation.** We use independently generated and selected alternative pseudo-GT targets to guide local teacher probing, and learn a shared correction mapping through restoration supervision, response-weighted consistency, and input content constraints. We evaluate PSR over 20 independent rounds of supervision regeneration and teacher retraining for each of Nano-Banana-2 and GPT-Image-2. For Nano-Banana-2, PSR reduces cross-batch perceptual differences and pixel-level fluctuations by 54.99% and 56.96%, respectively, relative to the corresponding original teacher.
- **Lightweight image restoration through composite teacher distillation.** We distill the composite teacher comprising the restoration teacher and PSR into a student with only 0.289 M parameters. The student outperforms all evaluated baselines on all six no-reference image quality metrics while retaining stability close to that of the composite teacher. Compared with the best baseline result for each metric, the student reduces NIQE by 16.30% and increases MANIQA and CLIP-IQA by 14.33% and 14.49%, respectively.

2 Related Work

Real-world image restoration. BSRGAN (Zhang et al., 2021) and Real-ESRGAN (Wang et al., 2021) construct training data by modeling diverse degradations, but discrepancies between synthetic and real degradations limit model generalization. MIDD (Flepp et al., 2024) provides real-world denoising data from multiple sensors, although acquiring paired images is constrained by capture and alignment requirements. Noise2Self (Batson & Royer, 2019) learns denoising from noisy observations under assumptions of conditionally independent noise and unbiased observations. AP-BSN (Lee et al., 2022) combines asymmetric pixel-shuffle downsampling with blind-spot networks to mitigate the effects of spatially correlated noise. We pair real degraded inputs with generated targets to study cross-scene image restoration without real clean supervision.

Content fidelity and stability of generative supervision. DiffBIR (Lin et al., 2024) combines degradation removal and information regeneration with region-adaptive guidance. GGT-100K (Kong et al., 2026) incorporates human review into training data construction. LIQE (Zhang et al., 2023) estimates perceptual quality but cannot directly establish content correspondence between inputs and targets. Noise2Same (Xie et al., 2020) combines reconstruction and invariance terms, while R-Drop (Wu et al., 2021) uses bidirectional KL divergence to constrain predictions

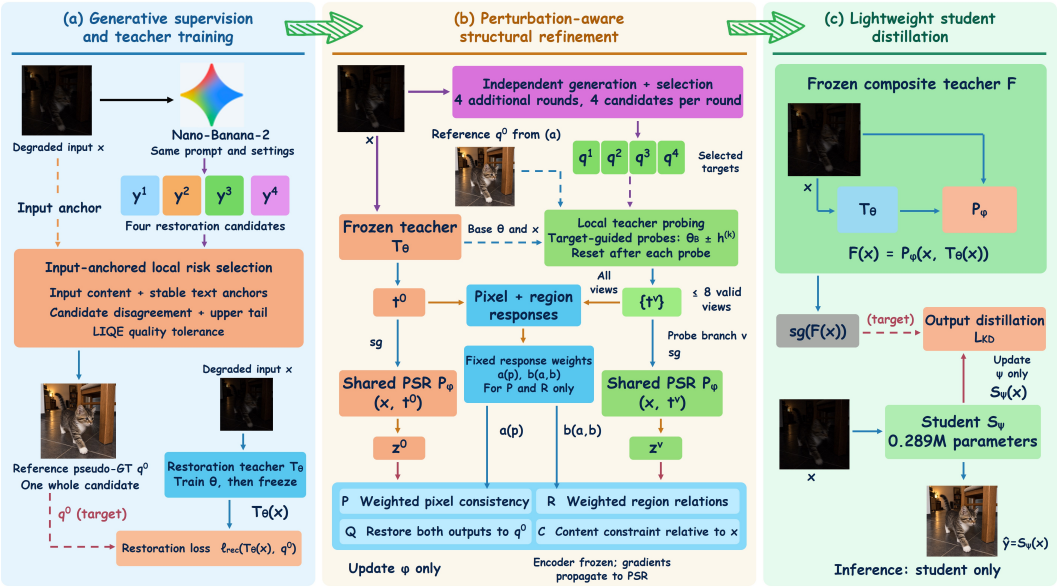

Figure 1: Overview of the proposed AIGC-supervised image restoration framework. (a) Input-anchored pseudo-target selection and teacher training. (b) Local teacher probing and response-weighted PSR training. (c) Distillation from the frozen composite teacher to a lightweight student, which is used alone at inference.

under different dropout samples. We introduce input-anchored selection and PSR-based teacher correction to reduce the restoration model's sensitivity to variations in generative supervision.

Lightweight image restoration and knowledge distillation. NAFNet (Chen et al., 2022) builds an efficient restoration network using SimpleGate and simplified channel attention. AT-BSN (Chen et al., 2024b) obtains a lightweight denoising network through multi-teacher distillation. DCKD (Zhou et al., 2025) adapts the distillation process using dynamic contrastive constraints and distribution mapping. We train PSR after freezing the restoration teacher, then freeze the composite mapping formed by the teacher and PSR and train a compact student through output distillation.

3 Method

We develop an image restoration framework integrating generative supervision, teacher correction, and knowledge distillation. Figure 1 provides an overview of the framework.

3.1 Dataset

XSMax contains 9,900 real-world RGB images at a resolution of 1024×1024 pixels, spanning 8 coarse scene categories and 66 fine-grained scene categories. It includes mixed degradations such as noise, blur, underexposure, and artifacts. Coarse categories describe environment types, while fine-grained categories identify subjects within each environment. The training, validation, and test splits are disjoint at both category levels. The training set contains 6,000 images from 4 coarse and 40 fine-grained categories; the validation set contains 1,650 images from 2 coarse and 11 fine-grained categories; and the test set contains 2,250 images from 2 coarse and 15 fine-grained categories. Figure 2 shows the scene distribution and split assignments.

3.2 Generative Supervision Construction

Candidate generation and input-anchored local content constraints. For each training input x_i, Nano-Banana-2 generates $K = 4$ candidates, $\mathcal{Y}_i = \{y_i^{(j)}\}_{j=1}^{4}$, using the same prompt and generation settings. The prompt requests quality improvement while preserving object identity, text,

and spatial layout. Candidate files and identifiers are fixed before selection. Since overall perceptual quality may conceal local content changes, we combine input-anchored deviation, local candidate disagreement, and perceptual quality to select one complete candidate as the pseudo-GT target, denoted by $q_i^{(0)}$.

We extract $W = 49$ aligned, overlapping windows of size 256×256 from each input and its candidates. Each window is represented by a unit-norm feature $e_w(\cdot)$, with distance $d_{e,w}(u, v) = \frac{1}{2}\|e_w(u) - e_w(v)\|_2$. The local input-anchored deviation is defined as

$$a_{jw} = \max\left\{ d_{\mathrm{LF},w}(x, y^{(j)}), \rho_w^{\mathrm{sel}} d_{e,w}(x, y^{(j)}) \right\}, \tag{1}$$

where $d_{\mathrm{LF},w}$ denotes the mean absolute difference after low-pass filtering and downsampling. Using Gaussian smoothing G_s, we quantify input stability as

$$u_w^{\mathrm{sel}} = \underset{s\in\{0.6,1.2\}}{\mathrm{median}}\, d_{e,w}(x, G_s(x)), \qquad \rho_w^{\mathrm{sel}} = \exp\left(-\frac{u_w^{\mathrm{sel}}}{s_\rho^{\mathrm{sel}}}\right), \tag{2}$$

where the scale s_ρ^{sel} is estimated from training inputs, and ρ_w^{sel} measures representation stability. Stable text anchors require matching recognition results across the input and its two smoothed views, with a recognition confidence of at least 0.90 in each view. Altered or missing text, or insufficient recognition confidence in a candidate, contributes to its violation count h_j. Text that cannot be reliably recognized in the input is not used to assess candidate correctness. Further details are provided in Appendix B.1.

Local disagreement and target selection. We measure disagreement with the other candidates as $c_{jw} = \mathrm{median}_{l\neq j}\, d_{e,w}(y^{(j)}, y^{(l)})$ and combine it with input anchoring. To avoid diluting local anomalies, $\mathrm{Tail}_{0.10}$ averages the five highest window scores for each score type:

$$A_j = \mathrm{Tail}_{0.10}(\{a_{jw}\}), \qquad C_j = \mathrm{Tail}_{0.10}(\{c_{jw}\}), \qquad R_j = \max\left\{ \frac{A_j}{\tau_A}, \frac{C_j}{\tau_C} \right\} \tag{3}$$

The scales τ_A and τ_C are the 90th percentiles of the corresponding training scores. The frozen LIQE model (Zhang et al., 2023) provides a quality score $s_j = (Q_j - 1)/4$, where $Q_j \in [1, 5]$. All four candidates are evaluated at the same 20 fixed crop locations. Let $\mathcal{V}$ contain candidates that are successfully decoded, have the correct dimensions, and have valid scores for all required criteria. We define

$$\begin{aligned} \mathcal{F} &= \{ j \in \mathcal{V} : h_j = 0,\ A_j/\tau_A \le 1,\ C_j/\tau_C \le 1 \}, \\ \mathcal{B}_{\delta_{\mathrm{sel}}} &= \left\{ j \in \mathcal{F} : s_j \ge \max_{l\in\mathcal{V}} s_l - \delta_{\mathrm{sel}} \right\}, \end{aligned} \tag{4}$$

where $\delta_{\mathrm{sel}} = 0.025$. When $\mathcal{B}_{\delta_{\mathrm{sel}}}$ is nonempty, we select $j^\star = \arg\min_{j\in\mathcal{B}_{\delta_{\mathrm{sel}}}} R_j$ and set $q^{(0)} = y^{(j^\star)}$. Fallback rules and the treatment of unknown disagreement are detailed in Appendix B.3.

3.3 Restoration Teacher Training

The selected targets and their corresponding inputs form $\mathcal{D}_{\mathrm{pair}} = \{(x_i, q_i^{(0)})\}_{i=1}^{N_{\mathrm{pair}}}$, where $N_{\mathrm{pair}} = 6{,}000$ when every input receives a target. We train the teacher using its restoration loss ℓ_{rec}:

$$\theta^\star = \arg\min_\theta \frac{1}{N_{\mathrm{pair}}} \sum_{i=1}^{N_{\mathrm{pair}}} \ell_{\mathrm{rec}}\left(T_\theta(x_i), q_i^{(0)} \right) \tag{5}$$

The reference pseudo-GT targets remain fixed during training, and identical spatial transformations are applied to each input–target pair. After training, we freeze the teacher and proceed to PSR correction.

3.4 Perturbation-aware Structural Refinement

Selected targets and local teacher probing. For each input x_i, we perform four additional independent rounds of generation and selection. Each round generates four candidates and selects one

target using the full selection rule, yielding $\{q_i^{(k)}\}_{k=1}^4$. These targets capture residual supervision variation after selection, rather than differences among rejected candidates. They are used only for local teacher probing; teacher training and PSR restoration supervision continue to use $q_i^{(0)}$.

The teacher operates in evaluation mode with fixed statistics. We restrict probing to the last nonlinear decoder block and the output head, collectively denoted by parameter block B. Let $d = HWC$ and $\ell(\theta, x, q) = d^{-1}\|T_\theta(x) - q\|_2^2$. We define

$$\begin{aligned} \widetilde{q}_i^{(k)} &= (1-\alpha)q_i^{(0)} + \alpha q_i^{(k)}, \\ g_i^{(k)} &= \nabla_{\theta_B}\ell\left(\theta, x_i, \widetilde{q}_i^{(k)}\right) - \nabla_{\theta_B}\ell\left(\theta, x_i, q_i^{(0)}\right) \end{aligned} \tag{6}$$

Let $r_B = \|\theta_B\|_2$, with the scale s_g determined from nonzero gradient-difference norms during training, and let $h_{\max} = 2\kappa r_B$. We calibrate and clip the perturbation as

$$\gamma = \frac{\kappa r_B}{s_g}, \qquad h_i^{(k)} = \gamma g_i^{(k)} \min\left\{1, \frac{h_{\max}}{\|\gamma g_i^{(k)}\|_2 + \varepsilon}\right\} \tag{7}$$

We then temporarily perturb the parameter block in the positive and negative directions to obtain

$$t_i^0 = T_\theta(x_i), \qquad t_i^{k,\sigma} = T_{\theta +_B \sigma h_i^{(k)}}(x_i), \qquad \sigma \in \{-1, +1\} \tag{8}$$

Here, $+_B$ indicates that only parameter block B is modified. Each probe starts from the original parameters without accumulating updates, yielding at most eight probe views. Probes with invalid calibration scales or outputs are discarded. If no valid views remain, we use the original teacher output and set the explicit consistency terms to zero.

Responses and shared correction. For the set of valid probe views $\mathcal{V}_i^{\text{probe}}$, we measure pixel responses and changes in inter-region relationships:

$$\begin{aligned} u_i(p) &= \frac{1}{|\mathcal{V}_i^{\text{probe}}|C} \sum_{v \in \mathcal{V}_i^{\text{probe}}} \sum_c \left(t_i^v(p,c) - t_i^0(p,c)\right)^2, \\ v_i(a,b) &= \frac{1}{|\mathcal{V}_i^{\text{probe}}|} \sum_{v \in \mathcal{V}_i^{\text{probe}}} \left(R_{ab}(t_i^v) - R_{ab}(t_i^0)\right)^2 \end{aligned} \tag{9}$$

Here, p and c denote pixel position and channel, respectively, and $R_{ab}(y) = f_a(y)^\top f_b(y)$ for $a < b$. The unit-norm region features f_a are obtained from a frozen encoder using pooling over an 8×8 grid. Fixed, bounded, nonnegative, and increasing mappings produce the weights $a_i(p) = \omega_u(u_i(p); \tau_u)$ and $b_i(a,b) = \omega_R(v_i(a,b); \tau_R)$. These weights remain fixed during training, imposing stronger consistency constraints on sensitive regions and inter-region relationships. When calibration is invalid, the corresponding weights are set to zero.

PSR applies a residual correction:

$$\begin{aligned} P_\phi(x,t) &= t + \delta_{\text{psr}} \tanh\left(R_\phi([x,t])\right), \\ z_i^0 &= P_\phi(x_i, t_i^0), \qquad z_i^v = P_\phi(x_i, t_i^v) \end{aligned} \tag{10}$$

Here, $[x,t]$ denotes the channel-wise concatenation of the degraded input and the teacher output, and R_ϕ predicts a three-channel correction. The two branches share the PSR parameters.

Joint training objective. Restoration supervision uses the same reference pseudo-GT for both branches, while response-weighted consistency constrains differences in pixels and inter-region relationships:

$$\begin{aligned} \mathcal{L}_Q &= \frac{1}{2d}\left(\|z_i^0 - q_i^{(0)}\|_2^2 + \|z_i^v - q_i^{(0)}\|_2^2\right), \\ \mathcal{L}_P &= \frac{1}{d}\sum_{p,c} a_i(p)\left(z_i^0(p,c) - z_i^v(p,c)\right)^2, \\ \mathcal{L}_R &= \frac{1}{2016}\sum_{a<b} b_i(a,b)\left(R_{ab}(z_i^0) - R_{ab}(z_i^v)\right)^2 \end{aligned} \tag{11}$$

Table 1: No-reference image quality results on the XSMax test set. Ours (teacher) denotes the composite teacher comprising the restoration teacher and the full PSR module. Ours (student) denotes the distilled student with 0.289 M parameters. Bold indicates the best result in each column.

Method	NIQE↓	MUSIQ↑	MANIQA↑	AFINE-NR↓	CLIP-IQA↑	TOPIQ↑
Noise2Self (Batson & Royer, 2019)	7.7322	52.7987	0.5124	−0.6437	0.4336	0.4042
DA-CLIP (Luo et al., 2024)	7.4731	54.0164	0.5281	−0.6775	0.4554	0.4215
UniRestore (Chen et al., 2025b)	7.4056	54.9653	0.5291	−0.6855	0.4599	0.4296
AdaIR (Cui et al., 2025)	7.3547	55.8425	0.5301	−0.6944	0.4608	0.4353
R2R (Wang et al., 2026)	7.3189	56.7498	0.5342	−0.7039	0.4677	0.4379
GGT-100K (Kong et al., 2026)	7.2070	57.4774	0.5477	−0.7584	0.4769	0.4530
Ours (teacher)	**6.0058**	**63.8766**	**0.6288**	**−1.0180**	**0.5483**	**0.5046**
Ours (student)	6.0323	63.5715	0.6262	−1.0134	0.5460	0.5023

To limit excessive changes to the input content, we uniformly sample four of the 49 windows without replacement to form $\mathcal{J}$ and penalize feature deviations exceeding a tolerance ξ:

$$\mathcal{L}_C = \frac{1}{2|\mathcal{J}|} \sum_{w \in \mathcal{J}} \rho_{iw}^{\mathrm{psr}} \sum_{b \in \{0,v\}} \left[\|e_w(z_i^b) - e_w(x_i)\|_2^2 - \xi \right]_+ \tag{12}$$

Here, $[r]_+ = \max(r, 0)$. For $v_{iw} = \mathrm{median}_{s \in \{0.6, 1.2\}} \|e_w(x_i) - e_w(G_s(x_i))\|_2^2$, we set $\rho_{iw}^{\mathrm{psr}} = \exp(-v_{iw}/s_\rho^{\mathrm{psr}})$ and $\xi = Q_{0.90}(\{v_{iw}\}_{\mathrm{train}})$. The scale s_ρ^{psr} is the larger of the median training-window response and a numerical lower bound (Appendix B.5). The overall objective is

$$\mathcal{L}_{\mathrm{PSR}} = \mathcal{L}_Q + \lambda_P \mathcal{L}_P + \lambda_R \mathcal{L}_R + \lambda_C \mathcal{L}_C \tag{13}$$

Only ϕ is updated. The encoder weights remain frozen, while gradients from the corrected images are propagated through the encoder to PSR.

3.5 Knowledge Distillation for Lightweight Deployment

After training PSR, we freeze the composite mapping $F(x) = P_\phi(x, T_\theta(x))$ and transfer its restoration capabilities to a student network through output-based knowledge distillation:

$$\mathcal{L}_{\mathrm{KD}} = \frac{1}{N_{\mathrm{pair}} d} \sum_{i=1}^{N_{\mathrm{pair}}} \|S_\psi(x_i) - \mathrm{sg}\left(F(x_i)\right)\|_2^2 \tag{14}$$

Only the student parameters ψ are updated, where sg denotes the stop-gradient operator. At deployment, only the student is retained, producing the restored output $\hat{y} = S_\psi(x)$.

4 Experiments

4.1 Experimental Setup

We use the XSMax splits described in Section 3.1, with Nano-Banana-2 as the primary generative model. The composite teacher comprises the restoration teacher and the full PSR module, while the distilled student has 0.289 M parameters. All methods are evaluated on the same 2,250 test images using NIQE (Mittal et al., 2012), MUSIQ (Ke et al., 2021), MANIQA (Yang et al., 2022), AFINE-NR (Chen et al., 2025a), CLIP-IQA (Wang et al., 2023), and TOPIQ (Chen et al., 2024a).

4.2 Restoration Performance on Unseen Scenes

Table 1 shows that both the composite teacher and the student outperform all evaluated external baselines on all six metrics. Compared with the best baseline result for each metric, the student reduces NIQE by 16.30% and increases MUSIQ, MANIQA, CLIP-IQA, and TOPIQ by 10.60%, 14.33%, 14.49%, and 10.88%, respectively. AFINE-NR decreases by 0.2550 in absolute terms.

As shown in Figure 3, Noise2Self still exhibits noticeable graininess and blurred edges. Although AdaIR, R2R, and GGT-100K improve visibility, local lines and character boundaries remain indistinct. In comparison, our method more clearly reveals patterns in dark regions, tire markings, leaf

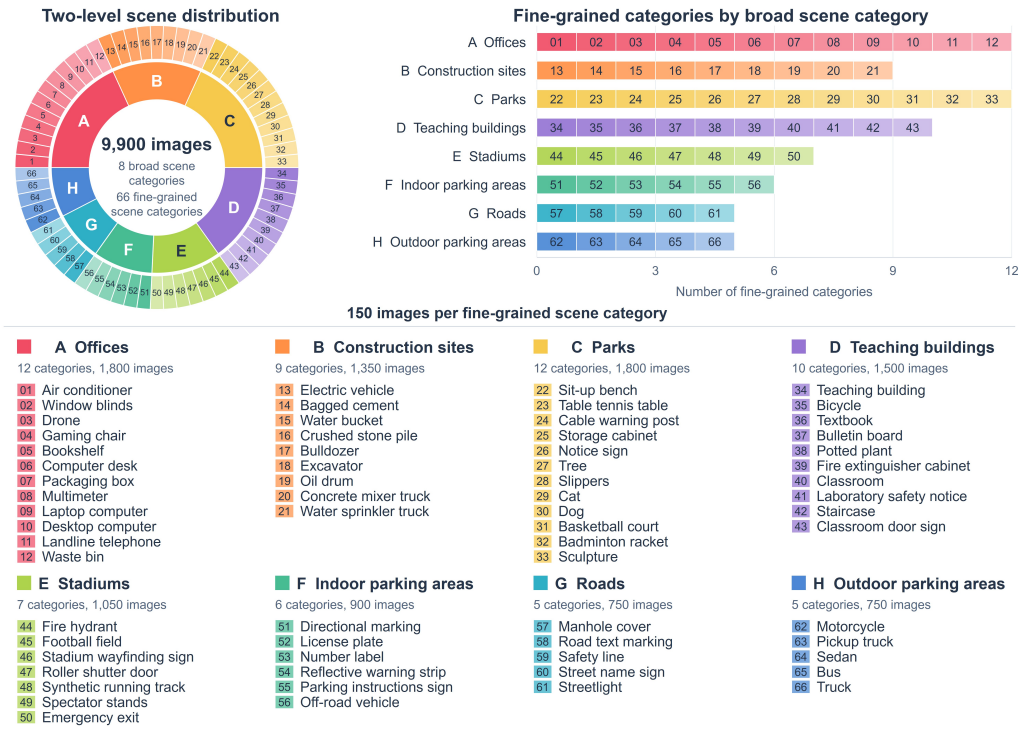

Figure 2: Scene distribution and data splits of XSMax: Training: A–C and E; validation: F–G; test: D and H.

Table 2: Ablation results for generative supervision selection. Bold indicates the best result in each column.

Selection configuration	NIQE ↓	MUSIQ ↑	MANIQA ↑	AFINE-NR ↓	CLIP-IQA ↑	TOPIQ ↑
Random-1	6.9701	58.6884	0.5637	−0.7942	0.4938	0.4604
Q-only	6.7933	59.3145	0.5701	−0.8025	0.5002	0.4625
Global	6.7042	59.7280	0.5797	−0.8284	0.5077	0.4686
Full − Anchor	6.7451	59.5092	0.5744	−0.8189	0.5037	0.4669
Full − Disagreement	6.6055	60.3443	0.5897	−0.8458	0.5144	0.4708
Full − Tail	6.5112	60.5433	0.5924	−0.8768	0.5186	0.4735
Full, $\delta_{\text{sel}} = 0$	6.3288	61.8372	0.6061	−0.9221	0.5299	0.4844
Full	**6.0323**	**63.5715**	**0.6262**	**−1.0134**	**0.5460**	**0.5023**

details, and character strokes in the examples shown, providing qualitative support for the results in Table 1.

4.3 Ablation Study of Generative Supervision Selection

We compare eight selection configurations under the same experimental protocol. All configurations are detailed in Appendix C.3.1. Table 2 shows that the full configuration achieves the best scores on all six metrics, reducing NIQE by 13.45% relative to random selection. Removing either input anchoring or candidate disagreement degrades performance on all six metrics. Compared with the variants using mean aggregation and zero quality tolerance, the full configuration reduces NIQE by 7.36% and 4.68%, respectively.



Figure 3: Qualitative comparison on real-world degraded images.

Table 3: Restoration stability under variations in generative supervision. Δ denotes the standard deviation of test-set mean quality scores over 20 rounds.

Source	Model	$\Delta_{\text{NIQE}}\downarrow$	$\Delta_{\text{MUSIQ}}\downarrow$	$\Delta_{\text{MANIQA}}\downarrow$	$\Delta_{\text{AFINE-NR}}\downarrow$	$\Delta_{\text{CLIP-IQA}}\downarrow$	$\Delta_{\text{TOPIQ}}\downarrow$	**PairLPIPS**$\downarrow$ ratio	**PixelSTD**$\downarrow$ ratio
Nano-Banana-2	Original teacher	0.1324	1.0142	0.0186	0.0477	0.0197	0.0113	1.0000	1.0000
	Teacher + full PSR	**0.0432**	**0.4797**	**0.0089**	**0.0225**	**0.0092**	**0.0047**	**0.4501**	**0.4304**
GPT-Image-2	Original teacher	0.1741	1.1455	0.0201	0.0535	0.0205	0.0124	1.0000	1.0000
	Teacher + full PSR	**0.0844**	**0.6012**	**0.0091**	**0.0267**	**0.0095**	**0.0053**	**0.4872**	**0.4798**

4.4 Restoration Stability under Variations in Generative Supervision

For each of Nano-Banana-2 and GPT-Image-2, we conduct 20 independent rounds of pseudo-GT generation, selection with the full configuration, and teacher training under fixed settings. For each generative source, we train a separate PSR module once using a reference teacher and its probe data, then freeze it and apply it to the 20 teachers from that source. Comparisons before and after correction use matched teacher checkpoints and identical test images. We report the standard deviation of per-round mean quality scores, together with PairLPIPS, based on LPIPS (Zhang et al., 2018), and PixelSTD. Complete definitions are provided in Appendix C.4.2.

Table 3 shows that PSR reduces all reported measures of variability. Relative to the corresponding original teacher, PairLPIPS and PixelSTD decrease by 54.99% and 56.96%, respectively, for Nano-Banana-2, and by 51.28% and 52.02%, respectively, for GPT-Image-2. These results indicate that PSR mitigates the effects of variations in pseudo-supervision on restoration outputs, improving stability across supervision generation rounds.

4.5 Ablation Study of PSR Components

Using Nano-Banana-2, we compare seven configurations with the same correction architecture, reference targets, training budget, and probe views. Here, Q, P, R, and C denote the four loss terms

Table 4: Ablation of PSR training objectives and response weighting with Nano-Banana-2 as the generative source.

Configuration	$\Delta_{\text{NIQE}}\downarrow$	$\Delta_{\text{MUSIQ}}\downarrow$	$\Delta_{\text{MANIQA}}\downarrow$	$\Delta_{\text{AFINE-NR}}\downarrow$	$\Delta_{\text{CLIP-IQA}}\downarrow$	$\Delta_{\text{TOPIQ}}\downarrow$	**PairLPIPS** $\downarrow$ ratio	**PixelSTD** $\downarrow$ ratio
Original teacher	0.1324	1.0142	0.0186	0.0477	0.0197	0.0113	1.0000	1.0000
Q only	0.1254	1.0002	0.0177	0.0435	0.0184	0.0102	0.9532	0.9334
$Q + C$	0.1124	0.9144	0.0159	0.0401	0.0167	0.0090	0.8124	0.8084
$Q + P + C$	0.0997	0.8332	0.0148	0.0362	0.0152	0.0081	0.7337	0.7258
$Q + P + R$	0.0812	0.7564	0.0134	0.0331	0.0137	0.0072	0.6859	0.6681
Fixed 0.5	0.0696	0.6131	0.0115	0.0293	0.0116	0.0059	0.5774	0.5518
Full PSR	**0.0432**	**0.4797**	**0.0089**	**0.0225**	**0.0092**	**0.0047**	**0.4501**	**0.4304**

Table 5: Cross-batch stability of the composite teacher (restoration teacher + full PSR) and distilled student using Nano-Banana-2.

Model	$\Delta_{\text{NIQE}}\downarrow$	$\Delta_{\text{MUSIQ}}\downarrow$	$\Delta_{\text{MANIQA}}\downarrow$	$\Delta_{\text{AFINE-NR}}\downarrow$	$\Delta_{\text{CLIP-IQA}}\downarrow$	$\Delta_{\text{TOPIQ}}\downarrow$	**PairLPIPS** $\downarrow$ ratio	**PixelSTD** $\downarrow$ ratio
Composite teacher	0.0432	0.4797	0.0089	0.0225	0.0092	0.0047	0.4501	0.4304
Distilled student	0.0433	0.4796	0.0087	0.0224	0.0093	0.0048	0.4502	0.4305

defined in Section 3.4. The fixed-weight variant retains all four terms and sets the response weights in P and R to 0.5. Complete configuration details are provided in Appendix C.5.1. Table 4 shows that full PSR achieves the best results on all eight measures. Using only Q reduces PairLPIPS and PixelSTD by just 4.68% and 6.66%, respectively, relative to the original teacher. Compared with the fixed-weight variant, full PSR further reduces these two measures by 22.05% and 22.00%, respectively.

4.6 Lightweight Knowledge Distillation from the Composite Teacher

For each of the 20 Nano-Banana-2 experimental rounds, we train one student from the corresponding frozen composite teacher. All students use the same architecture and training configuration, with the full protocol provided in Appendix C.6.1. Table 5 shows that the standard deviations of the quality scores differ by at most 0.0002 between the composite teacher and student. The student's PairLPIPS and PixelSTD ratios are 0.4502 and 0.4305, respectively, each only 0.0001 higher than the corresponding composite teacher ratio. These results indicate that output distillation can transfer the composite teacher's restoration mapping to the student while largely preserving its cross-batch stability.

5 Conclusion

To address real-world image restoration without real clean targets, we construct a dataset of captured images with diverse degradations and propose a framework that integrates generative supervision, PSR, and knowledge distillation. We select pseudo-GT targets using input-anchored local risk constraints and pair them with real degraded inputs to train a restoration teacher. To mitigate residual variations in generative supervision after selection, we freeze the teacher and introduce PSR, which uses alternative pseudo-GT targets to guide local teacher probing and learns a shared correction mapping through restoration supervision, response-weighted consistency, and input content constraints. The resulting composite teacher is then distilled into a student model. Experiments demonstrate that the proposed framework supports both restoration quality and output stability under variations in generative supervision, providing a practical approach to image restoration and lightweight deployment without real clean targets.

References

Joshua Batson and Loic Royer. Noise2self: Blind denoising by self-supervision. In *International conference on machine learning*, pp. 524–533. PMLR, 2019.

Chaofeng Chen, Jiadi Mo, Jingwen Hou, Haoning Wu, Liang Liao, Wenxiu Sun, Qiong Yan, and Weisi Lin. Topiq: A top-down approach from semantics to distortions for image quality assessment. *IEEE Transactions on Image Processing*, 33:2404–2418, 2024a.

Du Chen, Tianhe Wu, Kede Ma, and Lei Zhang. Toward generalized image quality assessment: Relaxing the perfect reference quality assumption. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12742–12752. IEEE, 2025a.

I Chen, Wei-Ting Chen, Yu-Wei Liu, Yuan-Chun Chiang, Sy-Yen Kuo, Ming-Hsuan Yang, et al. Unirestore: Unified perceptual and task-oriented image restoration model using diffusion prior. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 17969–17979, 2025b.

Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image restoration. In *European conference on computer vision*, pp. 17–33. Springer, 2022.

Shiyan Chen, Jiyuan Zhang, Zhaofei Yu, and Tiejun Huang. Exploring efficient asymmetric blind-spots for self-supervised denoising in real-world scenarios. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2814–2823. IEEE, 2024b.

Yuning Cui, Syed Waqas Zamir, Salman Khan, Alois Knoll, Mubarak Shah, and Fahad Khan. Adair: Adaptive all-in-one image restoration via frequency mining and modulation. In *International Conference on Learning Representations*, volume 2025, pp. 101306–101327, 2025.

Roman Flepp, Andrey Ignatov, Radu Timofte, and Luc Van Gool. Real-world mobile image denoising dataset with efficient baselines. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 22368–22377. IEEE, 2024.

Junjie Ke, Qifei Wang, Yilin Wang, Peyman Milanfar, and Feng Yang. Musiq: Multi-scale image quality transformer. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 5128–5137. IEEE, 2021.

Xiangtao Kong, Jixin Zhao, Lingchen Sun, Rongyuan Wu, and Lei Zhang. Ggt-100k: Generative ground truth for generalizable real-world image restoration. *arXiv preprint arXiv:2605.31039*, 2026.

Wooseok Lee, Sanghyun Son, and Kyoung Mu Lee. Ap-bsn: Self-supervised denoising for real-world images via asymmetric pd and blind-spot network. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 17704–17713. IEEE, 2022.

Xinqi Lin, Jingwen He, Ziyan Chen, Zhaoyang Lyu, Bo Dai, Fanghua Yu, Yu Qiao, Wanli Ouyang, and Chao Dong. Diffbir: Toward blind image restoration with generative diffusion prior. In *European conference on computer vision*, pp. 430–448. Springer, 2024.

Ziwei Luo, Fredrik K Gustafsson, Zheng Zhao, Jens Sjölund, and Thomas Schön. Controlling vision-language models for multi-task image restoration. In *International Conference on Learning Representations*, volume 2024, pp. 16226–16246, 2024.

Anish Mittal, Rajiv Soundararajan, and Alan C Bovik. Making a "completely blind" image quality analyzer. *IEEE Signal processing letters*, 20(3):209–212, 2012.

Chenxu Wang, Kai Zhang, and Jian Yang. Retrieve-to-restore: Efficient all-in-one image restoration with a retrieval-based degradation bank. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1277–1287, 2026.

Jianyi Wang, Kelvin CK Chan, and Chen Change Loy. Exploring clip for assessing the look and feel of images. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, pp. 2555–2563, 2023.

Xintao Wang, Liangbin Xie, Chao Dong, and Ying Shan. Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, pp. 1905–1914. IEEE, 2021.

Lijun Wu, Juntao Li, Yue Wang, Qi Meng, Tao Qin, Wei Chen, Min Zhang, Tie-Yan Liu, et al. R-drop: Regularized dropout for neural networks. *Advances in neural information processing systems*, 34:10890–10905, 2021.

Yaochen Xie, Zhengyang Wang, and Shuiwang Ji. Noise2same: Optimizing a self-supervised bound for image denoising. *Advances in neural information processing systems*, 33:20320–20330, 2020.

Sidi Yang, Tianhe Wu, Shuwei Shi, Shanshan Lao, Yuan Gong, Mingdeng Cao, Jiahao Wang, and Yujiu Yang. Maniqa: Multi-dimension attention network for no-reference image quality assessment. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1190–1199. IEEE, 2022.

Kai Zhang, Jingyun Liang, Luc Van Gool, and Radu Timofte. Designing a practical degradation model for deep blind image super-resolution. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 4771–4780. IEEE, 2021.

Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *2018 IEEE/CVF conference on computer vision and pattern recognition*, pp. 586–595. IEEE, 2018.

Weixia Zhang, Guangtao Zhai, Ying Wei, Xiaokang Yang, and Kede Ma. Blind image quality assessment via vision-language correspondence: A multitask learning perspective. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 14071–14081. IEEE, 2023.

Yunshuai Zhou, Junbo Qiao, Jincheng Liao, Wei Li, Simiao Li, Jiao Xie, Yunhang Shen, Jie Hu, and Shaohui Lin. Dynamic contrastive knowledge distillation for efficient image restoration. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 10861–10869, 2025.

A Use of a Large Language Model

In this manuscript, we used a large language model (LLM) solely for language editing to improve sentence clarity and correct grammatical errors. The LLM was not involved in the development of research ideas, methodology design, experimental implementation, or result analysis. All technical contributions and experimental results presented in this paper were independently produced and verified by the authors.

B Additional Method Details

B.1 Input Anchoring and Selection Calibration

Let j index candidates and w denote a local window. We extract 49 overlapping windows of size 256×256 pixels at corresponding spatial locations in the input and each candidate. Using a frozen encoder, we extract and pool features within each window, then normalize them to obtain a unit-norm representation $e_w(\cdot)$. The local feature distance is defined as

$$d_{e,w}(u, v) = \frac{1}{2} \left\| e_w(u) - e_w(v) \right\|_2 \tag{15}$$

To reduce the influence of unstable input details on candidate selection, we construct two input views using Gaussian smoothing with standard deviations of 0.6 and 1.2 pixels. We measure window sensitivity using the median feature distance between the original input and its smoothed views:

$$u_w^{\mathrm{sel}} = \operatorname*{median}_{s \in \{0.6, 1.2\}} d_{e,w}(x, G_s(x)), \tag{16}$$

$$\begin{aligned} \rho_w^{\mathrm{sel}} &= \exp\left(-\frac{u_w^{\mathrm{sel}}}{s_\rho^{\mathrm{sel}}} \right), \\ s_\rho^{\mathrm{sel}} &= \max\left\{ \operatorname*{median}_{i,w} u_{iw}^{\mathrm{sel}}, 10^{-6} \right\} \end{aligned} \tag{17}$$

Here, G_s denotes Gaussian smoothing with standard deviation s, and the scale s_ρ^{sel} is estimated from training inputs. This weight reflects the stability of the input representation under mild smoothing; it does not represent the probability that the content is correct.

We further normalize RGB pixel values to $[0, 1]$ and compute the mean absolute difference between corresponding windows after low-pass filtering and downsampling, denoted by $d_{\mathrm{LF},w}$. The local deviation of a candidate from the input anchor is defined as

$$a_{jw} = \max\left\{ d_{\mathrm{LF},w}(x, y^{(j)}), \rho_w^{\mathrm{sel}} d_{e,w}(x, y^{(j)}) \right\} \tag{18}$$

The low-frequency term captures differences in color and coarse spatial layout, while the feature term captures shifts in local representations. For text and numbers, we also construct stable text anchors. A text region is used as a constraint if it can be spatially matched across the original input and both smoothed views, yields identical recognition results, and has a recognition confidence of at least 0.90 in all three views. For each such anchor, altered or missing text, or insufficient recognition confidence in the corresponding candidate region, counts as one violation. The total number of violations is denoted by h_j. Text that cannot be consistently recognized in the input is not used to judge whether characters in a candidate are correct.

B.2 Disagreement, Upper-Tail Aggregation, and Quality Scores

Local differences among candidates provide complementary information about the instability of generated outputs. For candidate j, we compute its feature distances to the other three candidates within corresponding windows and take the median:

$$c_{jw} = \operatorname*{median}_{l \neq j} d_{e,w}\left(y^{(j)}, y^{(l)}\right) \tag{19}$$

Since agreement among candidates does not imply content correctness, we use this disagreement term jointly with input-anchored constraints.

To prevent image-wide averaging from obscuring local anomalies, we aggregate window scores using the upper-tail mean. Sorting the scores in descending order as $v_{(1)} \geq \cdots \geq v_{(W)}$, we define

$$\mathrm{Tail}_\eta(v) = \frac{1}{m} \sum_{r=1}^{m} v_{(r)}, \qquad m = \lceil \eta W \rceil \tag{20}$$

We separately aggregate input-anchored deviations and disagreement among candidates:

$$\begin{aligned} A_j &= \mathrm{Tail}_\eta\left(\{a_{jw}\}_{w=1}^{W}\right), \\ C_j &= \mathrm{Tail}_\eta\left(\{c_{jw}\}_{w=1}^{W}\right) \end{aligned} \tag{21}$$

We set $W = 49$ and $\eta = 0.10$, so each aggregation averages the five highest window scores. These windows are selected separately for the two score types.

B.2.1 Scale Calibration and Target Selection

We calibrate the scales of input-anchored deviation and candidate disagreement using the score distributions of training candidates:

$$\begin{aligned} \tau_A &= \max\left\{ Q_{0.90}(\mathcal{A}_{\mathrm{tr}}), 10^{-6} \right\}, \\ \tau_C &= \max\left\{ Q_{0.90}(\mathcal{C}_{\mathrm{tr}}^{\mathrm{obs}}), 10^{-6} \right\} \end{aligned} \tag{22}$$

Here, $Q_{0.90}$ denotes the 90th percentile computed using a fixed linear interpolation rule, and $\mathcal{C}_{\mathrm{tr}}^{\mathrm{obs}}$ contains valid disagreement estimates.

The perceptual quality of each candidate is estimated by the frozen LIQE model (Zhang et al., 2023). Let $Q_j \in [1, 5]$ denote the expected quality rating under the model's predicted distribution over five quality levels. The normalized quality score is

$$s_j = \frac{Q_j - 1}{4} \tag{23}$$

For the four candidates associated with the same input, we evaluate quality at the same 20 fixed crop locations to reduce the influence of crop-location differences on the comparison.

B.3 Exceptional Cases in Selection

Let $\mathcal{V}$ denote the set of candidates that can be successfully decoded, have the correct dimensions, and yield valid scores for all required criteria. Define $\bar{c}_j = C_j/\tau_C$. When the feasible set under the proxy constraints is empty but valid candidates remain, we rank the valid candidates lexicographically by the text violation count h_j, proxy risk R_j, negative quality score $-s_j$, and fixed candidate identifier, in that order. We select the first-ranked candidate and mark it as a forced fallback. The fallback target retains the original training weight but is not considered to satisfy all content constraints. If no usable disagreement estimates are available in the training set, we do not compute τ_C and apply the unknown-disagreement rule throughout.

B.4 Calibration of Local Teacher Probing

Interpolated targets are used only to construct probing directions and do not replace the restoration supervision. We calibrate the perturbation magnitude using the scale s_g, derived from nonzero gradient-difference norms on the training calibration set. Let $r_B = \|\theta_B\|_2$ and $h_{\max} = 2\kappa r_B$. We then define

$$\begin{aligned} \gamma &= \frac{\kappa r_B}{s_g}, \\ h_i^{(k)} &= \gamma g_i^{(k)} \min\left\{1, \frac{h_{\max}}{\|\gamma g_i^{(k)}\|_2 + \varepsilon}\right\} \end{aligned} \tag{24}$$

B.5 Input Content Weights and Tolerance

The two consistency losses are normalized by the number of image elements and the number of region pairs, respectively. To limit excessive changes to input content, we reuse the local window representations e_w from the frozen encoder and compute content weights based on the sensitivity of the input to mild Gaussian smoothing:

$$v_{iw} = \underset{s\in\{0.6,1.2\}}{\text{median}} \|e_w(x_i) - e_w(G_s(x_i))\|_2^2, \tag{25}$$

$$\begin{aligned} \rho_{iw}^{\text{psr}} &= \exp\left(-\frac{v_{iw}}{s_\rho^{\text{psr}}}\right), \\ \xi &= Q_{0.90}\left(\{v_{iw}\}_{\text{train}}\right) \end{aligned} \tag{26}$$

Here, s_ρ^{psr} is the larger of the median training window response and a numerical lower bound. At each training step, we uniformly sample four of the 49 windows without replacement to form the set $\mathcal{J}$. We then use Equation equation 12 to penalize feature deviations exceeding the tolerance ξ.

C Detailed Experimental Protocols and Analysis

C.1 Experimental Setup

C.1.1 Data and Model Configuration

We use the XSMax data splits described in Section 3.1. The dataset comprises images captured using an Apple iPhone XS Max, with 6,000 training images, 1,650 validation images, and 2,250 test images. Nano-Banana-2 serves as the primary generative model. The restoration teacher is trained using the selected pseudo-GT targets and subsequently combined with the full PSR module to form the composite teacher. The student is trained through knowledge distillation to match the outputs of the frozen composite teacher and contains 0.289 M parameters. In Table 1, Ours (teacher) denotes the composite teacher, and Ours (student) denotes the distilled student.

C.1.2 Image Quality Assessment Metrics

We evaluate restoration results using NIQE (Mittal et al., 2012), MUSIQ (Ke et al., 2021), MANIQA (Yang et al., 2022), AFINE-NR (Chen et al., 2025a), CLIP-IQA (Wang et al., 2023), and TOPIQ (Chen et al., 2024a). Lower NIQE and AFINE-NR values indicate better quality, whereas

higher values are better for the other four metrics. We additionally use PairLPIPS and PixelSTD to evaluate stability. All methods are evaluated on the same test images.

C.2 Restoration Performance on Unseen Scenes

C.2.1 Overall Quantitative Comparison

Table 1 shows that both the composite teacher and the distilled student outperform all evaluated external baselines across the six no-reference image quality metrics. With only 0.289 M parameters, the student achieves a NIQE score of 6.0323, MUSIQ of 63.5715, MANIQA of 0.6262, AFINE-NR of -1.0134, CLIP-IQA of 0.5460, and TOPIQ of 0.5023. Compared with the best-performing external baseline, the student reduces NIQE by 16.30% and improves MUSIQ, MANIQA, CLIP-IQA, and TOPIQ by 10.60%, 14.33%, 14.49%, and 10.88%, respectively. AFINE-NR decreases from -0.7584 to -1.0134, corresponding to an absolute improvement of 0.2550. These results indicate that the proposed framework can transfer the restoration capability learned from generative supervision to a compact student, improving restoration quality on images captured by the same device in previously unseen scenes.

C.2.2 Qualitative Evaluation

As shown in Figure 3, Noise2Self outputs still exhibit pronounced graininess and blurred edges. Although AdaIR, R2R, and GGT-100K improve visibility in some regions, pattern lines and character boundaries remain indistinct. In comparison, our method produces clearer local structures in these examples. Fine lines and contours in dark patterns are more distinct, and character boundaries on the tire markings are clearer. Improved visibility in dark vegetation regions makes leaf contours, overlapping structures, and some leaf veins easier to discern. Text strokes also appear clearer. These visual comparisons demonstrate improved restoration quality on the illustrated real-world examples with mixed degradations, particularly in the clarity of local structures and details, providing qualitative support for the no-reference quality assessment in Table 1.

C.3 Ablation Study of Generative Supervision Selection

C.3.1 Comparison Settings

All configurations use the same training inputs and the same set of four initial candidates for each input. Each configuration selects one complete candidate image as the reference pseudo-GT. For each configuration, we then retrain the restoration teacher, train the corresponding PSR module, and perform knowledge distillation to evaluate the effect of the selection strategy on the complete restoration pipeline. All configurations share the same subsequent training procedure and budget. PSR probing targets are constructed according to the selection rules of the corresponding configuration.

We compare the following eight configurations:

- **Random-1:** Select one valid candidate using a predefined random selection list.
- **Q-only:** Select the candidate with the highest LIQE quality score.
- **Global:** Jointly consider quality and deviation from the input using $s_j - \lambda D_j$, where s_j is the normalized LIQE score and D_j is the mean input-anchored deviation across windows. The coefficient λ is calibrated using the interquartile ranges of training candidate scores to match their scales.
- **Full − Anchor:** Remove input-anchored scoring and text constraints while retaining candidate disagreement and quality-based selection.
- **Full − Disagreement:** Remove candidate disagreement and select targets based on input-anchored risk and perceptual quality.
- **Full − Tail:** Replace the upper-tail mean with the mean over all windows and re-estimate the corresponding scales on the training set.
- **Full, $\delta_{\mathrm{sel}} = 0$:** Set the quality tolerance to zero while keeping all other selection settings unchanged.

- **Full:** Apply the complete selection rule.

Whenever a component is removed, its associated feasibility constraints, ranking terms, and fallback rules are also removed. For example, after removing candidate disagreement, the proxy risk becomes $R_j = A_j/\tau_A$. The placeholder treatment for unknown disagreement is not used as a substitute for this ablation.

C.3.2 Results and Analysis

Table 2 shows that the full selection strategy achieves the best reported results across all six quality metrics. Compared with Random-1 and Q-only, Full reduces NIQE by 13.45% and 11.20%, respectively. These results suggest that considering local content risk in addition to perceptual quality improves restoration. Full also outperforms the globally weighted scoring strategy (Global) on the reported metrics. Removing either input-anchored constraints or candidate disagreement degrades performance across all six metrics. Compared with mean aggregation and zero quality tolerance, Full reduces NIQE by 7.36% and 4.68%, respectively, supporting the effectiveness of aggregating high-risk regions and allowing a quality tolerance. These results indicate that the full selection strategy improves the restoration quality of the final student trained using generative supervision.

C.4 Cross-Batch Stability of PSR

C.4.1 Experimental Protocol

We use Nano-Banana-2 and GPT-Image-2 as generative models and conduct 20 independent rounds of pseudo-GT generation, selection, and teacher training for each model. The XSMax training inputs, generation settings, full selection rule, and teacher training configuration remain fixed across rounds. For each generative model, we train a separate PSR module once using a reference teacher and its corresponding probing data. We then freeze the trained PSR module and apply it to the 20 restoration teachers trained with supervision from that generative model. To obtain paired comparisons, we evaluate each teacher before and after PSR correction using the same teacher checkpoint, the same 2,250 test images, and identical evaluation settings.

C.4.2 Stability Metrics

Let $y_{i,r}^{(A)}$ denote the restoration output of method A for test image i in experimental round r, where $M = 2{,}250$ and $R = 20$. For each quality metric m, we first compute the mean score over the test set in each round, then calculate the sample standard deviation of these scores across the 20 rounds:

$$\begin{aligned} m_r^{(A)} &= \frac{1}{M}\sum_{i=1}^{M} m\left(y_{i,r}^{(A)}\right), \\ \Delta_m(A) &= \sqrt{\frac{1}{R-1}\sum_{r=1}^{R}\left(m_r^{(A)} - \bar{m}^{(A)}\right)^2} \end{aligned} \tag{27}$$

Here, $\bar{m}^{(A)} = R^{-1}\sum_{r=1}^{R} m_r^{(A)}$ is the mean score across rounds.

We further use LPIPS (Zhang et al., 2018) to measure perceptual differences between restoration outputs for the same input across experimental rounds, and summarize these differences as PairLPIPS. For each test image, we evaluate all 190 unordered pairs among its 20 outputs and then average over the test set:

$$\mathrm{PairLPIPS}(A) = \frac{2}{MR(R-1)}\sum_{i=1}^{M}\sum_{1\le r<s\le R}\mathrm{LPIPS}\left(y_{i,r}^{(A)}, y_{i,s}^{(A)}\right) \tag{28}$$

PixelSTD measures variation in pixel values across rounds:

$$\mathrm{PixelSTD}(A) = \frac{1}{Md}\sum_{i=1}^{M}\sum_{p=1}^{d}\sqrt{\frac{1}{R-1}\sum_{r=1}^{R}\left(y_{i,r}^{(A)}(p) - \bar{y}_i^{(A)}(p)\right)^2} \tag{29}$$

Here, $d = HWC$, p indexes all spatial and channel elements, and $\bar{y}_i^{(A)}$ is the mean output across rounds.

To facilitate comparison, we use the original teacher T from the same generative source as the reference and compute

$$r_{\mathrm{LPIPS}}(A) = \frac{\mathrm{PairLPIPS}(A)}{\mathrm{PairLPIPS}(T)}, \qquad r_{\mathrm{STD}}(A) = \frac{\mathrm{PixelSTD}(A)}{\mathrm{PixelSTD}(T)} \tag{30}$$

Both ratios equal 1 for the original teacher. Values below 1 indicate reduced output variability. The corresponding relative reduction is $(1 - r) \times 100\%$, where r denotes the respective ratio.

C.4.3 Results and Analysis

Table 3 shows that PSR reduces the standard deviations of all six quality scores for both generative models. For Nano-Banana-2, the standard deviation of NIQE decreases from 0.1324 to 0.0432, a reduction of 67.37%, while that of MUSIQ decreases from 1.0142 to 0.4797, a reduction of 52.70%. The largest reduction among the remaining four metrics is 58.41%. For GPT-Image-2, all six standard deviations decrease, with a maximum reduction of 57.26%.

For Nano-Banana-2, the PairLPIPS and PixelSTD ratios are 0.4501 and 0.4304, corresponding to reductions of 54.99% and 56.96% relative to the original teacher. For GPT-Image-2, the corresponding ratios are 0.4872 and 0.4798, representing reductions of 51.28% and 52.02%, respectively. These results indicate that PSR attenuates the propagation of variation in pseudo-supervision to restoration outputs, improving the stability of the restoration system across independent rounds of supervision generation.

C.5 Ablation Study of PSR Components

C.5.1 Comparison Settings

Using Nano-Banana-2 as the primary generative model, we compare seven configurations to examine the effects of correction training objectives and response weighting on stability. All trainable correction modules share the same architecture, reference restoration targets, training budget, and probing views.

Here, Q denotes restoration supervision, which encourages the corrected outputs of the original and probing branches to match the same reference pseudo-GT. P denotes response-weighted pixel consistency, R denotes response-weighted regional relation consistency, and C denotes the input content constraint. The fixed-weight configuration sets all response weights in P and R to 0.5 while keeping the other loss components and training settings unchanged.

The Q-only configuration uses standard residual correction. The $Q + C$, $Q + P + C$, and $Q + P + R$ configurations remove explicit consistency, regional relation consistency, and the input content constraint, respectively. Both the fixed-weight configuration and full PSR retain all four loss terms.

C.5.2 Results and Analysis

Table 4 shows that full PSR achieves the best results for the standard deviations of all six quality scores and both output stability metrics. Using restoration supervision alone reduces PairLPIPS and PixelSTD by only 4.68% and 6.66%, respectively, relative to the original teacher, indicating limited stability gains. Adding explicit consistency constraints to restoration supervision and the input content constraint reduces the PairLPIPS and PixelSTD ratios from 0.8124 and 0.8084 to 0.4501 and 0.4304, respectively. Removing either regional relation consistency or the input content constraint increases both measures of output variability. Compared with fixing the response weights at 0.5, full PSR further reduces PairLPIPS and PixelSTD by 22.05% and 22.00%, respectively.

C.6 Distilling the Composite Teacher into a Lightweight Student

C.6.1 Experimental Setup

To evaluate whether the corrected restoration mapping and its stability can be transferred to a lightweight network, we train 20 students using the 20 frozen composite teachers obtained with the primary generative model, Nano-Banana-2, with one student per teacher. Each composite teacher consists of the corresponding restoration teacher and the same frozen full PSR module. Each student receives only degraded images and learns to match the restoration outputs of its corresponding composite teacher using the output distillation objective in Section 3.5. All students share the same network architecture and training configuration. The restoration teachers and PSR remain frozen throughout distillation, and only student parameters are updated. At deployment, only the student with 0.289 M parameters is retained. The teacher, PSR, generative model, and feature encoder are not used during student inference.

C.6.2 Retention of Restoration Quality and Stability

Table 5 shows that the student and composite teacher have similar standard deviations across all six quality scores, with absolute differences no greater than 0.0002. The student's PairLPIPS and PixelSTD ratios are 0.4502 and 0.4305, respectively, each exceeding the corresponding composite teacher ratio by only 0.0001. Relative to the original teacher from the same generative source, these values correspond to reductions of 54.98% and 56.95% in the two output variability measures, closely matching the reductions achieved by the composite teacher. Together, Tables 1 and 5 indicate that output-based knowledge distillation can transfer the composite teacher's restoration mapping to a student with 0.289 M parameters while largely preserving its stability across independent generation batches.