COUPLING-AWARE COMPOSITIONAL RESTORATION: ZERO-SHOT, TRAINING-FREE REMOVAL OF COMPOSITE ADVERSE WEATHER

Anonymous authors
Paper under double-blind review

ABSTRACT

Real-world adverse weather rarely appears in isolation: low light, haze, rain, and snow frequently co-occur and interact, producing *composite* degradations that are far harder to remove than any single effect. State-of-the-art methods for composite weather are fully supervised and must be trained on a combinatorially large set of degradation combinations, while existing training-free restorers built on diffusion priors are designed for a single, known degradation and break down once degradations are stacked. We present CACR, a *zero-shot, training-free* framework for composite adverse-weather restoration. CACR detects which base weathers are present, derives a physically consistent removal order, and applies image-space physical operators in a *coupling-aware* manner, re-estimating each operator's parameters from the partially restored image rather than from the raw input. No restoration network is trained and no paired restoration data is used; an optional few-shot probe only sharpens weather detection. We show that single-degradation zero-shot restorers can fall below the degraded input on stacked weather, that the coupling, not the individual operators, drives our improvement, and that diffusion-inversion editing is bounded by a reconstruction ceiling that motivates restoring in image space. On the CDD-11 composite-weather benchmark, CACR provides the first zero-shot, training-free reference point across all eleven classes, reported in PSNR, SSIM, and LPIPS, with a clear characterisation of the remaining gap to supervised state of the art.

1 INTRODUCTION

Images captured under adverse weather are a persistent obstacle for downstream computer-vision systems such as autonomous driving, outdoor surveillance, and aerial imaging. In practice, weather rarely manifests as a single, isolated effect: a night-time driving scene may simultaneously suffer from low illumination, haze, and rain, and these factors interact rather than simply add up. Such *composite* degradations are markedly harder to remove than any of their constituents in isolation, because each degradation alters the image statistics on which the removal of the others depends.

An extensive literature base addresses *single* known degradation, with dedicated models for deraining, dehazing, desnowing, or low-light enhancement. To avoid maintaining one model per degradation, all-in-one approaches learn a single network that handles several degradation *types*, either through contrastive degradation encoders (Li et al., 2022), learnable prompts (Potlapalli et al., 2023), or degradation-aware vision-language features (Luo et al., 2024). These methods, however, are typically trained and evaluated on one degradation at a time. Composite degradation has been targeted more directly by recent supervised frameworks such as OneRestore (Guo et al., 2024) and AllRestorer (Mao et al., 2024), which set the state of the art on the CDD-11 composite-weather benchmark (Guo et al., 2024). Their strength, though, comes at a cost: they require large-scale supervised training over many degradation combinations, whose number grows combinatorially with the set of base weathers (Cao et al., 2024).

A complementary, training-free line of work exploits pretrained diffusion models (Ho et al., 2020; Song et al., 2020) as image priors and casts restoration as inference in noise space. Posterior-sampling methods (Kawar et al., 2022; Chung et al., 2022) and diffusion-inversion methods (Chi-

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.

haoui et al., 2024; Chihaoui & Favaro, 2025) recover clean images without any task-specific training, and are appealing precisely because they need no paired data. Yet these methods are designed and validated for *single*, often linear or parametric, degradations; their underlying assumptions (a known forward operator, or a single off-distribution mode in the inverted noise) do not hold once several weathers are stacked. More recent unified efforts move toward mixtures of tasks (Li et al., 2025) or remove composite degradations step by step (Cao et al., 2024), but the latter still relies on *trained* multi-task base models. To the best of our knowledge, no training-free method has been demonstrated for composite adverse-weather restoration, and the CDD-11 benchmark remains occupied entirely by supervised models.

This gap motivates two questions that we study in this work. First, *how far can a training-free method go on composite weather?* Second, *why do existing zero-shot restorers fail when degradations are stacked?* We find that the failure is structural rather than incidental. Stacked weather degradations are *coupled*: the atmospheric light used by a dehazing model, for instance, is only meaningful once the scene illumination has been corrected, while brightening a hazy image first amplifies the haze. Consequently, applying single-degradation operators independently, or in an arbitrary fixed order, removes one artifact while injecting another. We further show, through a round-trip analysis, that diffusion-inversion editing on a generic prior is upper-bounded by a reconstruction ceiling that by itself precludes competitive restoration of these scenes, which motivates performing restoration in image space.

Building on these observations, we propose CACR, a *zero-shot, training-free* framework for composite adverse-weather restoration. Given a degraded image, CACR (i) detects which base weathers are present using a frozen vision backbone augmented with physical cues, (ii) peels them in a physically consistent order (illumination, then scattering, then particles), and (iii) applies image-space physical operators (illumination correction, depth-guided dark-channel dehazing, and targeted streak removal) in a *coupling-aware* manner, conditioning the order and parameters of each stage on the partially restored image and on the other detected degradations. Restoration trains no network, uses no paired data, and relies only on four global hyperparameters.

Our contributions are:

- We introduce CACR, to our knowledge the first *zero-shot, training-free* framework for *composite* adverse-weather restoration, providing the first such reference point on CDD-11, where existing zero-shot restorers fall *below* the degraded input.
- We identify and exploit the *coupling* between stacked degradations at *both* detection and restoration: haze is assessed only after deflaking and brightening, atmospheric light is re-estimated after illumination correction, and depth-guided dehazing is gated on the absence of low-light. Ablations show this coupling, not the individual operators, drives the improvement, distinguishing CACR from a naive cascade of classical filters.
- We combine frozen, off-the-shelf priors into an image-space operator suite (depth-guided transmission for dehazing and targeted masked de-streaking) that escapes the diffusion-inversion reconstruction ceiling we analyze.
- We evaluate CACR on CDD-11 against zero-shot baselines (diffusion-inversion and single-operator methods) in PSNR, SSIM, and LPIPS, winning on all composite classes, and report an honest gap to supervised state-of-the-art models.

2 Related Work

2.1 Single-Degradation and All-in-One Restoration

Classical image restoration develops a dedicated model for each degradation, such as deraining (Fu et al., 2017), dehazing (Li et al., 2017; Qin et al., 2020), desnowing (Liu et al., 2018), or low-light enhancement (Wei et al., 2018; Guo et al., 2020). Restoration under multiple adverse weathers has further been tackled with transformer backbones and feature-disentangling designs (Valanarasu et al., 2022; Zamir et al., 2022; Zhu et al., 2023). To reduce the cost of maintaining many specialised networks, all-in-one restoration trains a single model that handles several degradation *types*. AirNet learns degradation representations through contrastive learning (Li et al., 2022), PromptIR conditions a shared backbone on learnable degradation prompts (Potlapalli et al., 2023), and DA-CLIP

adapts a vision-language model to predict degradation-aware features that guide restoration (Luo et al., 2024). Recent surveys document the rapid growth of this paradigm (Jiang et al., 2025). These methods are trained and evaluated predominantly on *one* degradation at a time and do not model the interaction between co-occurring weathers.

2.2 Composite-Degradation Restoration

A smaller but growing body of work tackles *composite* degradations directly. OneRestore introduces the CDD-11 benchmark and a cross-attention framework driven by scene descriptors (Guo et al., 2024), while AllRestorer models the relationships among all present degradations in a single transformer block (Mao et al., 2024). Self-Improved Privilege Learning further boosts such models by reusing their own outputs as test-time guidance (Wu et al., 2025). These approaches define the state of the art on CDD-11, but they are fully supervised and require training over many degradation combinations, whose number grows combinatorially with the set of base weathers. Chain-of-Restoration mitigates this burden by training only on degradation bases and removing composite degradations step by step (Cao et al., 2024); it nevertheless still depends on *trained* multi-task base models. In contrast, CACR introduces no learned restoration network and requires no paired training data, providing the first zero-shot, training-free entry on this benchmark.

2.3 Training-Free Restoration with Diffusion Priors

Denoising diffusion models (Ho et al., 2020; Song et al., 2020) have become powerful generic image priors, enabling restoration without task-specific training. Posterior-sampling methods such as DDRM (Kawar et al., 2022), DPS (Chung et al., 2022) and generative diffusion priors (Fei et al., 2023) solve inverse problems by guiding the reverse diffusion process with a known forward operator, while inversion-based methods such as BIRD (Chihaoui et al., 2024) and Invert2Restore (Chihaoui & Favaro, 2025) optimise or rectify the inverted noise of a degraded image. LD-RPS extends zero-shot restoration toward mixed tasks via latent-diffusion posterior sampling (Li et al., 2025). Despite their generality, these methods are designed and validated for a single degradation, typically assuming a known linear operator or a single off-distribution mode in the inverted noise; both assumptions are violated when several weathers are stacked.

2.4 Physical Priors and Degradation Detection

Physically grounded operators remain strong, training-free tools for individual weather effects: the dark channel prior models atmospheric scattering for dehazing (He et al., 2010), and Retinex theory underpins illumination correction for low light (Land & McCann, 1971). We further ground dehazing in scene geometry with a frozen monocular depth estimator (Yang et al., 2024), since the atmospheric transmission decays with distance. Separately, large pretrained encoders provide zero- or few-shot degradation cues: CLIP offers open-vocabulary recognition (Radford et al., 2021), DINOv2 yields robust self-supervised features (Oquab et al., 2023), and DA-CLIP explicitly predicts degradation embeddings (Luo et al., 2024). CACR reuses such frozen components but *couples* them: operator order and parameters follow the detected weathers and are re-estimated as restoration proceeds.

3 Method

3.1 Composite Degradation as a Stack

A clean scene J observed under adverse weather is corrupted not by a single effect but by an ordered *composition* of base degradations. Let $g_w(\cdot\,;\theta_w)$ denote the image-formation operator of a base weather $w \in \mathcal{B} = \{\mathsf{low}, \mathsf{haze}, \mathsf{rain}, \mathsf{snow}\}$ with physical parameters θ_w (illumination gain, atmospheric light and transmission, streak layer). A composite image with a present set $\mathcal{W} \subseteq \mathcal{B}$, $|\mathcal{W}| = K$, is generated by stacking the corresponding operators,

$$I \;=\; \big(g_{w_K} \circ \cdots \circ g_{w_1}\big)(J), \qquad w_1, \ldots, w_K \in \mathcal{W}. \tag{1}$$

Crucially, these operators *do not commute*: haze applied to a dimmed scene is not the haze of a bright scene, and brightening a hazy image amplifies the veil. Hence a composite is fundamentally harder

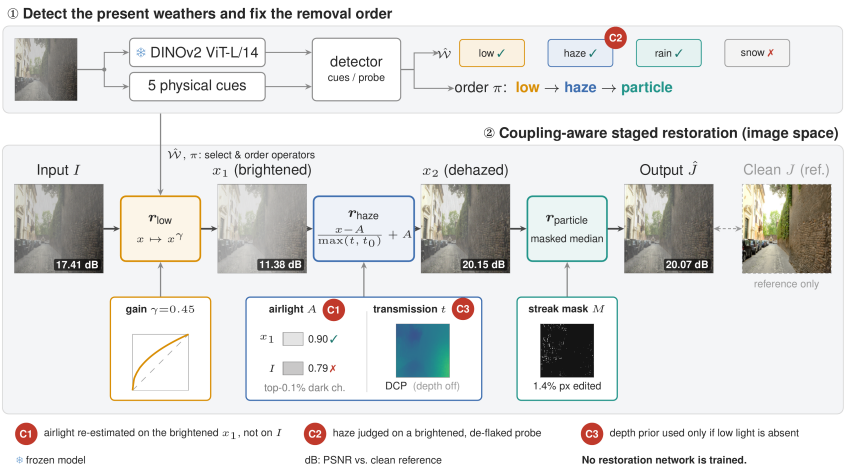

Figure 1: Encoder with five physical cues detects which base weathers are present; the composite is then inverted as an ordered sequence of *coupling-aware* physical operators (illumination → de-haze → de-particle), each re-estimating its parameters from the partially restored image. PSNR rises monotonically toward the clean reference. The three couplings (C1–C3) are detailed in Sec. 3.3. No restoration network is trained.

than the sum of its parts, and its inverse cannot be obtained by applying single-degradation methods independently. CACR instead inverts the stack as an ordered sequence of *coupled* single-degradation removals,

$$\hat{J} \;=\; \big(r_{w_1} \circ \cdots \circ r_{w_K}\big)(I), \qquad \hat{\theta}_{w_i} \;=\; \Theta_{w_i}\big(x_{i-1},\, \hat{\mathcal{W}}\big), \tag{2}$$

where r_w is the physical inverse operator for w, x_{i-1} is the image after the previous $i-1$ stages ($x_0 = I$), and the operator parameters $\hat{\theta}_{w_i}$ are re-estimated from the *current* (partially restored) image rather than the raw input. The pipeline (Figure 1) has three steps: *detect* $\mathcal{W}$ (Sec. 3.2), *order* the removal (Sec. 3.3), and *peel* each base weather with a coupling-aware physical operator (Sec. 3.4). All backbones (detection encoder, depth network) are frozen.

3.2 Weather Detection

We estimate $\hat{\mathcal{W}}$ from a descriptor that concatenates a frozen self-supervised embedding (Oquab et al., 2023) with five physical cues chosen to be discriminative where semantic features are ambiguous: dark-channel mean (haze, snow), high-frequency energy (rain, snow), mean luminance (low light), saturation and contrast (haze). A per-weather logistic probe maps this descriptor to presence probabilities. The cues are decisive for the hardest confusion, snow vs. haze: both raise the dark channel, but snow carries high high-frequency energy and haze does not. The detector has two modes: a *label-free* mode that thresholds the physical cues directly, and an optional probe fit on ∼20–30 labelled images per class. The probe affects only which operators are selected, never the restoration itself.

3.3 Coupling-Aware Staged Decomposition

The core of CACR is the decomposition of a composite into an ordered sequence of single-degradation problems. We remove degradations *coarse-to-fine*: global illumination first, then the volumetric scattering (haze), then the sparse high-frequency layer (rain/snow). This order follows the physics of Eq. (1): illumination and airlight are global and set the context for the finer layers, and is realised by the permutation π : low → haze → particle restricted to the detected set. Three couplings make the staging more than a cascade of filters:

(C1) Airlight after illumination. In the scattering model $I = J\,t + A(1-t)$, the airlight A is the brightness of the haze layer and is *underestimated* on a dimmed scene. We therefore correct illumination first and re-estimate A on the brightened image x_1, not on I.

Algorithm 1 CACR: coupling-aware staged restoration

Input: degraded image I; frozen encoder ϕ, depth net $\mathcal{D}$; params θ
Output: restored image $\hat{J}$

$\hat{\mathcal{W}} \leftarrow \textsc{Detect}(I, \phi)$ {semantic + physical cues}
order $\hat{\mathcal{W}}$ as low $\rightarrow$ haze $\rightarrow$ particle; $x \leftarrow I$
if low $\in \hat{\mathcal{W}}$ **then**
  $x \leftarrow x^{\gamma}$ {illumination}
end if
if haze $\in \hat{\mathcal{W}}$ **then**
  $A \leftarrow \textsc{Airlight}(x)$ {C1: on brightened x}
  if low $\notin \hat{\mathcal{W}}$ **then**
    $t \leftarrow \exp(-\beta\, \mathcal{D}(x))$ {C3}
  else
    $t \leftarrow \mathrm{DCP}(x, A, \omega)$
  end if
  $x \leftarrow (x - A)/\max(t, t_0) + A$
end if
if rain $\in \hat{\mathcal{W}}$ **or** snow $\in \hat{\mathcal{W}}$ **then**
  $x \leftarrow \textsc{MaskedDeStreak}(x, k)$
end if
return x

(C2) Haze assessed after de-flaking and brightening. White snow inflates the dark-channel cue and masquerades as haze, while low light suppresses it. The haze decision is therefore taken on a probe that is de-flaked (if rain/snow present) and brightened (if low light present), preventing a false dehaze that would corrupt a snowy scene.

(C3) Depth gated on low light. Depth-guided transmission (Sec. 3.4) needs a reliable monocular depth estimate, which collapses on dark scenes; it is therefore used only when low light is absent, with a dark-channel fallback otherwise.

Algorithm 1 summarises the procedure. As a worked example, a low+haze+rain image is detected as {low, haze, rain}, ordered low $\rightarrow$ haze $\rightarrow$ rain, then peeled: gamma illumination correction, dehazing with airlight re-estimated on the brightened image (C1), and targeted streak removal, each consuming the output of the previous stage.

3.4 Physical Inverse Operators

Illumination r_{low}. Gamma mapping x^{γ}, $\gamma < 1$, inverts global dimming.

Dehazing r_{haze}. We invert the scattering model, $\hat{J} = (x - A)/\max(t, t_0) + A$. The airlight A is the mean colour of the brightest 0.1% dark-channel pixels of the brightened image (C1). The transmission t uses either the dark-channel prior (He et al., 2010), $t = 1 - \omega \min_c \min_{y \in \Omega}\big(x_c(y)/A_c\big)$ with guided-filter (He et al., 2012) refinement, or, when low light is absent (C3), a *depth-guided* transmission $t(x) = \exp(-\beta\, d(x))$ where d is the normalised scene distance from a frozen monocular depth network (Yang et al., 2024). The latter grounds t in scene geometry rather than a heuristic prior, improving haze-dominant scenes.

Rain/snow removal r_{particle}. Streaks and flakes are bright, thin, high-frequency structures. We form a luminance detail layer $D = L - \mathrm{blur}(L)$, flag streak pixels $M = [\, D > k\, \sigma_D \,]$, and replace *only* the flagged pixels with a local median, leaving the background intact. This targeted operator removes particles without the blurring of a global filter, so it does no harm on near-clean inputs.

4 Experiments

We organise the evaluation around four questions. **(Q1)** Do existing zero-shot restorers actually help when degradations are *stacked*? **(Q2)** Does CACR improve over the degraded input across *every*

Method	Composite (7)			All (11)	
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑
Degraded input	15.42	0.685	0.331	16.69	0.725
DCP (He et al., 2010)	12.95	0.559	0.373	14.15	0.616
Gamma	13.55	0.647	0.405	14.02	0.688
BIRD (Chihaoui et al., 2024)	14.57	0.493	0.632	15.26	0.531
I2R (Chihaoui & Favaro, 2025)	13.42	0.459	0.730	13.28	0.463
LD-RPS (Li et al., 2025)	15.07	0.608	0.358	16.01	0.644
CACR (Ours)	**17.47**	**0.702**	**0.304**	**18.50**	**0.747**

Table 1: **Composite-weather restoration on CDD-11** (training-free methods). *Composite*: mean of the seven multi-degradation classes; *All*: all eleven. Every diffusion baseline scores *below* the input on composites; CACR improves all three metrics.

Setting	low+haze	Comp. mean
Coupling OFF	16.37	17.66
Coupling ON (Ours)	**17.20**	**18.01**
Det.: physical cues (label-free)	–	17.1
Det.: learned probe (Ours)	–	**17.5**
De-streak: global bilateral	rain 21.2 / snow 21.8	
De-streak: masked (Ours)	**rain 23.1 / snow 23.3**	

Table 2: **Ablations** (PSNR, dB). *Top*: coupling, with ground-truth detection and identical operators. *Middle*: detector, operators fixed. *Bottom*: targeted vs. global de-streaking.

composite class? **(Q3)** Is the gain due to the *coupling*, or merely to the classical operators? **(Q4)** Does a method designed for composites still generalise to *real*, single-weather data? Sec. 4.2–4.5 answer these in turn.

4.1 Experimental Setup

Benchmarks. Our primary benchmark is CDD-11 (Guo et al., 2024), the standard *composite* adverse-weather dataset: four single degradations (low light, haze, rain, snow), five double and two triple combinations (11 classes). We evaluate 100 images per class from the test split. To probe generalisation beyond synthetic composites, we additionally evaluate on six *real or independent* single-weather benchmarks used by recent all-in-one work (Mao et al., 2024): LOL-v2 real and synthetic (Yang et al., 2021), NH-HAZE (Ancuti et al., 2020), SPA-Data (Wang et al., 2019), and Snow100K / Snow100K-L (Liu et al., 2018).

Metrics and protocol. We report PSNR, SSIM and LPIPS against the clean reference at 256×256. LPIPS matters here: a diffusion prior can produce a plausible but *different* image, which LPIPS penalises. All external datasets are evaluated on their *full* test split; n is stated per row.

Implementation. Detection uses a frozen DINOv2 (Oquab et al., 2023) ViT-L/14 embedding plus five physical cues; depth-guided dehazing uses a frozen Depth-Anything-V2 (Yang et al., 2024). Restoration uses four global hyperparameters $(\gamma, \omega, \beta, k)$.

Baselines. We compare against training-free methods only, since that is our operating regime: classical single operators (DCP (He et al., 2010), gamma correction) and three diffusion-based zero-shot restorers: BIRD (Chihaoui et al., 2024) (no blind weather operator; we report its prior-only round-trip), our re-implementation of Invert2Restore (Chihaoui & Favaro, 2025), and LD-RPS (Li et al., 2025), a recent latent-diffusion recurrent posterior sampler. Supervised models trained on CDD-11 are *not* comparable to a training-free method and are discussed separately in Sec. 4.7.

4.2 Main Comparison on Composite Weather

Table 1 reports the headline result, and answers Q1 and Q2.

Q1: zero-shot diffusion restorers fall *below* the degraded input. Every diffusion baseline loses to simply doing nothing on composites: BIRD reaches 14.57 dB, I2R 13.42 dB and LD-RPS 15.07 dB, against a 15.42 dB input. The perceptual metrics are more damning still. LPIPS rises from 0.331 (input) to 0.632 (BIRD) and 0.730 (I2R), confirming that these methods do not merely fail to remove the degradation, they *substitute content*. LD-RPS is the strongest of the three and the closest to the input, yet it still degrades PSNR (-0.35 dB), SSIM (-0.077) and LPIPS ($+0.027$). Diffusion restoration seeks an image that is likely under the prior and consistent with a *known* degradation; with stacked, non-commuting weather the consistency term is mis-specified and the prior dominates, yielding clean-looking but different images. Classical single operators fail for the complementary reason: each addresses one weather and corrupts the rest, so DCP and gamma both fall ~ 2 dB below the input.

Method	low	haze	rain	snow	l+h	l+r	l+s	h+r	h+s	l+h+r	l+h+s
Degraded input	12.5	16.0	**23.1**	**23.9**	15.7	15.9	14.3	15.3	15.3	16.2	15.2
BIRD (Chihaoui et al., 2024)	13.2	15.1	18.2	19.3	14.6	14.8	14.2	14.4	14.9	14.8	14.4
I2R (Chihaoui & Favaro, 2025)	12.6	12.2	13.3	13.9	13.8	12.9	13.6	14.0	12.0	14.2	13.5
LD-RPS (Li et al., 2025)	13.1	15.6	19.7	20.1	16.3	16.3	14.7	14.4	14.4	15.0	14.2
CACR (Ours)	**18.7**	**19.4**	21.8	20.9	**16.8**	**18.5**	**16.7**	**19.2**	**18.4**	**17.1**	**15.7**

Table 3: **Per-class PSNR (dB) on CDD-11.** Abbreviations: l = low light, h = haze, r = rain, s = snow; e.g. l+h+r is the low + haze + rain triple. Best per class in **bold**.

Q2: CACR is the only training-free method that improves over the input. CACR reaches 17.47 dB composite ($+2.05$ dB over the input) and 18.50 dB over all 11 classes, while simultaneously achieving the best SSIM (0.702) and the best LPIPS (0.304). Improving PSNR and LPIPS jointly rules out gains from blurring or hallucination.

4.3 Per-Class Behaviour

Table 3 breaks the comparison down per class and reveals *where* each method succeeds or fails.

CACR improves every composite class. Gains range from $+0.4$ dB on the hardest triple (low+haze+snow) to $+3.9$ dB on haze+rain, and are largest exactly where the input is worst, namely severe low light ($12.5 \rightarrow 18.7$ dB, $+6.2$) and haze ($16.0 \rightarrow 19.4$). Restoration quality therefore scales with degradation severity, which is the regime that matters in practice.

Do-no-harm on near-clean inputs. On rain and snow the input is already at 23–24 dB. Here CACR intervenes only marginally (21.8 and 20.9 dB), because the masked de-streaking operator edits *only* flagged pixels. The contrast with the baselines is instructive: on the same two classes BIRD drops to 18.2/19.3 dB and I2R collapses to 13.3/13.9 dB. Explicit detection prevents this failure by construction.

4.4 Ablation Studies

We now isolate *which* component produces the gain (Q3). All ablations use ground-truth detection unless stated, so that detection noise does not confound the comparison. Results are in Table 2.

(a) The coupling, not the operators, drives the gain. Disabling the coupling (same operators, airlight estimated on the raw input, fixed order) costs $\mathbf{0.83}$ dB on the low+haze classes and 0.36 dB on the composite mean, while leaving single-degradation classes unchanged. This is the key control experiment: since the operator set is identical in both rows, the improvement can only come from re-estimating each operator's parameters on the partially restored image, separating CACR from a naive filter cascade.

(b) Detection quality, not restoration, bounds the remainder. Holding the operators fixed and varying only the detector, label-free physical cues reach 17.1 dB and the learned probe 17.5 dB. The gap is small, and both keep restoration fully zero-shot, since the probe only sharpens *which* operators fire. The remaining headroom is therefore a detection problem.

(c) Targeted operators beat global ones. Masked de-streaking outperforms global bilateral smoothing by $+1.9$ dB on rain and $+1.6$ dB on snow, because it edits only flagged pixels instead of blurring the background. Depth-guided transmission improves haze-dominant classes over the dark-channel prior (haze $19.5 \rightarrow 21.2$, haze+rain $19.3 \rightarrow 20.1$, haze+snow $18.5 \rightarrow 19.1$ dB) but *degrades* when low light co-occurs, since monocular depth is unreliable on dark scenes. This is the empirical justification for gating it on the absence of low light (coupling C3).

4.5 Generalisation to Real Single-Weather Data

A method tuned for synthetic composites could plausibly overfit to CDD-11's degradation model. To test this (Q4) we apply CACR *unchanged* to six independent benchmarks, including real-world haze and real rain (Table 4); qualitative examples appear in Fig. 2c.

Dataset	Weather	n	PSNR↑		SSIM↑		LPIPS↓	
			Input	CACR	Input	CACR	Input	CACR
LOL-v2 real (Yang et al., 2021)	low light	100	9.75	**18.00**	0.199	**0.765**	0.483	**0.164**
LOL-v2 syn. (Yang et al., 2021)	low light	100	10.59	**16.25**	0.406	**0.731**	0.420	**0.260**
NH-HAZE (Ancuti et al., 2020)	haze	55	11.74	**13.36**	0.444	**0.530**	0.555	**0.464**
Snow100K (Liu et al., 2018)	snow	500	19.58	**19.89**	0.740	**0.747**	0.226	**0.215**
Snow100K-L (Liu et al., 2018)	snow	16801	15.41	**15.69**	0.453	**0.462**	0.469	**0.463**
SPA-Data (Wang et al., 2019)	real rain	6385	34.23	34.08	0.919	**0.920**	0.169	**0.166**

Table 4: **Generalisation to independent single-weather benchmarks.** CACR is applied *unchanged*, with no per-dataset retuning, on the full test split of each set (n images). Better value in **bold**.



(a) Triple-degradation classes. (b) Double-degradation classes. (c) Independent benchmarks.

Figure 2: **Qualitative results.** (a,b) Staged restoration on CDD-11: columns are the input, each executed operator (+illumination, +dehaze, +de-particle), and the ground truth; per-panel PSNR rises *monotonically*, so no stage undoes a previous one. In (b) only detected operators fire, so rows follow different stage sequences. (c) LOL-v2 real, NH-HAZE, SPA-Data and Snow100K: input, CACR, ground truth, with no per-dataset tuning.

Large gains where degradation is severe. Low light is the strongest case: $+8.25$ dB on LOL-v2 real and $+5.66$ dB on LOL-v2 synthetic, with LPIPS reduced roughly threefold ($0.483 \rightarrow 0.164$). On NH-HAZE, which contains real non-homogeneous haze from a distribution never seen in CDD-11, CACR improves all three metrics ($+1.62$ dB). The physical operators transfer because they encode image formation, not dataset statistics.

Do-no-harm confirmed at scale. SPA-Data contains 6,385 real rainy images whose inputs already sit at 34.2 dB. CACR is essentially neutral in PSNR (-0.15 dB) yet still improves SSIM and LPIPS: faced with an almost-clean image, it declines to intervene rather than over-processing. The two snow benchmarks (17,301 images combined) show small but consistent gains on all three metrics.

Scope. These benchmarks cover the four base weathers our operators model. Degradations whose formation model differs fundamentally, for example raindrop occlusion where the underlying scene is locally *absent* rather than attenuated, fall outside the current operator set; extending the framework with an inpainting-style operator under the same coupling rules is natural future work.

4.6 Qualitative Results

Figure 2 shows the staged inversion on triple (2a) and double (2b) CDD-11 classes, where PSNR rises monotonically across stages, the visual signature of the coupling, and on real data for all four weathers (2c). Throughout, CACR removes the veil, brightens, and suppresses streaks while pre-

serving scene structure; the diffusion baselines instead alter content, consistent with their high LPIPS in Table 1.

4.7 Discussion

Relation to supervised models. A model trained on CDD-11 such as OneRestore (Guo et al., 2024) reaches $\sim$28 dB, well above our 17.5 dB. We view these as answering different questions. Supervised composite restoration requires paired data for a combinatorially growing set of degradation combinations, eleven classes here but 2^K in general, and its accuracy on unseen combinations is not guaranteed. CACR asks instead how far one can go with *no restoration training at all*, and shows that a coupling-aware inversion of physical operators recovers a substantial part of the gap while remaining applicable to any combination its operators cover.

5 Conclusion

We presented CACR, a zero-shot, training-free framework for *composite* adverse-weather restoration. Treating stacked degradations as a non-commuting composition, CACR detects the present weathers, orders their removal coarse-to-fine, and peels them with coupling-aware image-space operators. On CDD-11 it is the only training-free method that improves over the degraded input on every composite class, while diffusion baselines fall below it. Ablations attribute the gain to the coupling rather than the individual operators and locate the remaining headroom in detection.

Limitations. CACR remains well below supervised composite restoration (17.5 vs. $\sim$28 dB on CDD-11), and its gains depend on detection: a missed weather means its operator never runs, a false positive can degrade a clean input, and the strongest results use a few-shot probe, with smaller gains from purely label-free cues. The operator set and the fixed low$\rightarrow$haze$\rightarrow$particle order cover only the four CDD-11 base weathers; degradations with a different formation model (raindrop occlusion, blur, noise) are not handled, and the four global hyperparameters are shared across all images rather than adapted per scene. On near-clean rain and snow inputs the masked de-streaking still costs 1–3 dB PSNR, and the depth prior cannot be used on dark scenes. Finally, all composite results are on the synthetic CDD-11 at 256×256 and our real-data evaluation covers single weathers only, as real paired data with co-occurring weather is limited to few combinations.

Future work. The coupling-aware staged formulation is general and extensible: additional frozen physical or learned operators (for blur, noise, or JPEG) can be inserted as new stages under the same detect-order-peel schedule. Stronger few-shot detection and geometry-aware operators are promising routes toward supervised performance without training a restoration network.

AI use statement

We used generative AI tools (Claude Opus 5 and Claude Opus 5.5, Anthropic) for formatting, readability editing, and help with writing. All content was reviewed by the authors, who take full responsibility for this work.

Reproducibility statement

CACR trains no restoration network; all results are determined by frozen public backbones (DINOv2 ViT-L/14, Depth-Anything-V2) and fixed settings shared across datasets: $\gamma = 0.45$, $\omega = 0.8$, $t_0 = 0.1$, $k = 2.5$, with inputs resized to 256×256. The detection probe is fit on 30 CDD-11 training images per class. All benchmarks are public (Sec. 4.1), and code will be released upon acceptance.

References

Codruta O Ancuti, Cosmin Ancuti, and Radu Timofte. Nh-haze: An image dehazing benchmark with non-homogeneous hazy and haze-free images. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pp. 444–445, 2020.

Jin Cao, Deyu Meng, and Xiangyong Cao. Chain-of-restoration: Multi-task image restoration models are zero-shot step-by-step universal image restorers. *arXiv preprint arXiv:2410.08688*, 2024.

Hamadi Chihaoui and Paolo Favaro. Invert2restore: Zero-shot degradation-blind image restoration. *arXiv preprint arXiv:2503.21486*, 2025.

Hamadi Chihaoui, Abdelhak Lemkhenter, and Paolo Favaro. Blind image restoration via fast diffusion inversion. *Advances in Neural Information Processing Systems*, 37:34513–34532, 2024.

Hyungjin Chung, Jeongsol Kim, Michael T Mccann, Marc L Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. *arXiv preprint arXiv:2209.14687*, 2022.

Ben Fei, Zhaoyang Lyu, Liang Pan, Junzhe Zhang, Weidong Yang, Tianyue Luo, Bo Zhang, and Bo Dai. Generative diffusion prior for unified image restoration and enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 9935–9946, 2023.

Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John Paisley. Removing rain from single images via a deep detail network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3855–3863, 2017.

Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 1780–1789, 2020.

Yu Guo, Yuan Gao, Yuxu Lu, Huilin Zhu, Ryan Wen Liu, and Shengfeng He. Onerestore: A universal restoration framework for composite degradation. In *European conference on computer vision*, pp. 255–272. Springer, 2024.

Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2341–2353, 2010.

Kaiming He, Jian Sun, and Xiaoou Tang. Guided image filtering. *IEEE transactions on pattern analysis and machine intelligence*, 35(6):1397–1409, 2012.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Junjun Jiang, Zengyuan Zuo, Gang Wu, Kui Jiang, and Xianming Liu. A survey on all-in-one image restoration: Taxonomy, evaluation and future trends. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025.

Bahjat Kawar, Michael Elad, Stefano Ermon, and Jiaming Song. Denoising diffusion restoration models. *Advances in neural information processing systems*, 35:23593–23606, 2022.

Edwin H Land and John J McCann. Lightness and retinex theory. *Journal of the Optical society of America*, 61(1):1–11, 1971.

Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. Aod-net: All-in-one dehazing network. In *Proceedings of the IEEE international conference on computer vision*, pp. 4770–4778, 2017.

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

Huaqiu Li, Yong Wang, Tongwen Huang, Hailang Huang, Haoqian Wang, and Xiangxiang Chu. Ld-rps: Zero-shot unified image restoration via latent diffusion recurrent posterior sampling. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 13684–13694, 2025.

Yun-Fu Liu, Da-Wei Jaw, Shih-Chia Huang, and Jenq-Neng Hwang. Desnownet: Context-aware deep network for snow removal. *IEEE Transactions on Image Processing*, 27(6):3064–3073, 2018.

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.

Jiawei Mao, Yu Yang, Xuesong Yin, Ling Shao, and Hao Tang. Allrestorer: All-in-one transformer for image restoration under composite degradations. *arXiv preprint arXiv:2411.10708*, 2024.

Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, et al. Dinov2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193*, 2023.

Vaishnav Potlapalli, Syed Waqas Zamir, Salman H Khan, and Fahad Shahbaz Khan. Promptir: Prompting for all-in-one image restoration. *Advances in neural information processing systems*, 36:71275–71293, 2023.

Xu Qin, Zhilin Wang, Yuanchao Bai, Xiaodong Xie, and Huizhu Jia. Ffa-net: Feature fusion attention network for single image dehazing. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pp. 11908–11915, 2020.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, pp. 8748–8763. PmLR, 2021.

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

Jeya Maria Jose Valanarasu, Rajeev Yasarla, and Vishal M Patel. Transweather: Transformer-based restoration of images degraded by adverse weather conditions. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 2353–2363, 2022.

Tianyu Wang, Xin Yang, Ke Xu, Shaozhe Chen, Qiang Zhang, and Rynson WH Lau. Spatial attentive single-image deraining with a high quality real rain dataset. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 12270–12279, 2019.

Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*, 2018.

Gang Wu, Junjun Jiang, Kui Jiang, and Xianming Liu. Boosting all-in-one image restoration via self-improved privilege learning. *arXiv preprint arXiv:2505.24207*, 2025.

Lihe Yang, Bingyi Kang, Zilong Huang, Zhen Zhao, Xiaogang Xu, Jiashi Feng, and Hengshuang Zhao. Depth anything v2. *Advances in Neural Information Processing Systems*, 37:21875–21911, 2024.

Wenhan Yang, Wenjing Wang, Haofeng Huang, Shiqi Wang, and Jiaying Liu. Sparse gradient regularized deep retinex network for robust low-light image enhancement. *IEEE Transactions on Image Processing*, 30:2072–2086, 2021.

Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5728–5739, 2022.

Yurui Zhu, Tianyu Wang, Xueyang Fu, Xuanyu Yang, Xin Guo, Jifeng Dai, Yu Qiao, and Xiaowei Hu. Learning weather-general and weather-specific features for image restoration under multiple adverse weather conditions. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 21747–21758, 2023.