LUCE: Constrained Curve-Domain Guidance for Training-Free Low-Light Enhancement with Hue-Preserving Decoupling

Yijie Wei[1,2], Weiran Li[1,2], Yeqiang Liu[1,2], Mina Han[1,2], Xue Liu[1], and Zhenbo Li[1,2,3,4,5,6(✉)]

[1] College of Information and Electrical Engineering, China Agricultural University, Beijing, China
{yjwei,vranlee,yeqiangliu,nanako,liusnow,lizb}@cau.edu.cn
[2] National Innovation Center for Digital Fishery, Ministry of Agriculture and Rural Affairs, Beijing, China
[3] Key Laboratory of Agricultural Informatization Standardization, Ministry of Agriculture and Rural Affairs, Beijing, China
[4] Key Laboratory of Smart Farming Technologies for Aquatic Animal and Livestock, Ministry of Agriculture and Rural Affairs, Beijing, China
[5] State Key Laboratory of Efficient Utilization of Agricultural Water Resources, Beijing, China
[6] National Digital Agricultural Product Circulation (Supply Chain and Logistics) Innovation Sub-Center, Ministry of Agriculture and Rural Affairs, Beijing, China

Abstract. Low-light image enhancement (LLIE) on pretrained diffusion models is commonly achieved through weighted pixel-space guidance losses, whose competing gradients often destabilize sampling and induce color distortion. We present LUCE, a training-free framework that reformulates sampler-side guidance as a constrained curve-domain luminance control problem. Instead of directly optimizing pixel intensities, LUCE projects luminance guidance into a low-dimensional monotonic curve space and enforces a unified luminance target via a single scalar energy defined on a piecewise-linear LUT. This constrained parameterization restricts guidance to a globally consistent exposure trajectory and removes the need for balancing multiple heterogeneous losses. To preserve chromatic fidelity, we introduce a hue-preserving chromatic decoupling mechanism that realizes luminance edits through per-pixel scalar gains while maintaining RGB direction. Implemented as a gradient-based correction within a frozen DDPM sampler, LUCE requires no retraining or architectural modification. Experiments on paired and unpaired benchmarks demonstrate stable exposure adjustment and improved color fidelity over prior training-free diffusion baselines.

Keywords: Constrained curve-domain luminance control · Hue-preserving chromatic decoupling · Gradient-based correction

✉ Corresponding author.

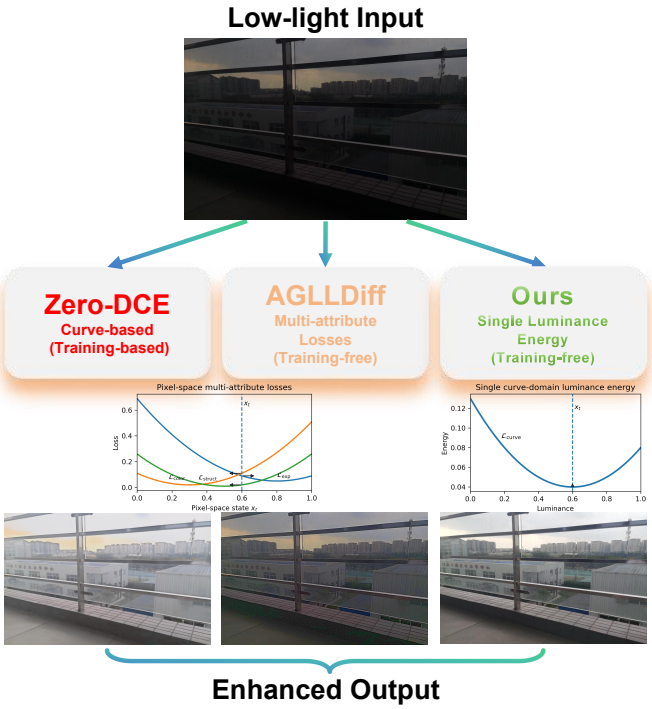

Fig. 1: Motivation of LUCE. Given the same low-light input, Zero-DCE (curve-based, training-based) and AGLLDiff (training-free diffusion with multiple pixel-space attribute losses) represent two dominant LLIE paradigms. The insets visualize their guidance formulation: AGLLDiff aggregates several heterogeneous pixel-space losses with competing gradients at state x_t, whereas LUCE replaces them with a single curve-domain luminance energy. With chromatic decoupling, this single-energy guidance produces a well-exposed result with natural colors and preserved structure.

1 Introduction

Images captured under low illumination often suffer from low visibility, noise, and color distortion, degrading both human perception and downstream vision performance. Low-light image enhancement (LLIE) seeks to restore visibility and contrast while preserving natural color and structure, and has long been studied in computational photography and low-level vision [17,41]. Classical methods based on histogram equalization or Retinex-style illumination estimation can improve brightness but tend to amplify noise, create halos, and struggle in complex real-world scenes with mixed degradations [7,29].

Deep learning has advanced LLIE by learning data-driven illumination priors. Supervised methods train convolutional or transformer-based networks on paired datasets such as LOL [39,45], achieving strong enhancement but at the cost of labor-intensive data collection and limited generalization across cameras and

capture conditions. Zero-reference approaches such as Zero-DCE [6] estimate image-specific tone curves from non-reference losses, offering an interpretable curve formulation, yet still require offline training and can fail under extremely dark or spatially heterogeneous low light.

Flow-based models like LLFlow [37] and diffusion-based methods trained in RAW, wavelet, or Retinex-inspired domains [9, 15, 18, 19, 30, 33] achieve strong results but require costly retraining and task-specific architectures. In contrast, recent training-free diffusion frameworks perform LLIE directly on top of pre-trained diffusion backbones by adding auxiliary objectives during sampling. Works such as AGLLDiff [24] and PGDiff [43] define exposure-, structure-, and color-related objectives in pixel space and steer sampling via weighted combinations of their gradients. While effective, this paradigm suffers from three drawbacks: (i) gradient competition, where heterogeneous objectives pull in different directions and require delicate weight tuning; (ii) luminance–chroma entanglement, since RGB-space losses often distort hue while brightening; and (iii) implicit brightness evolution, lacking an explicit global model of a smooth, monotonic exposure trajectory. Fig. 1 illustrates these issues on a challenging scene, and contrasts pixel-space multi-attribute losses with our single curve-domain luminance energy.

Diffusion sampling naturally admits gradient-based guidance, where objective gradients are injected into each update step to steer the trajectory. This motivates a constrained view of sampler-side LLIE: instead of combining several pixel-space losses on x_t, we treat exposure adjustment as constrained luminance control, project luminance cues into a single scalar energy in a 1D curve domain, and decouple luminance edits from chroma via a hue-preserving transform.

LUCE instantiates this principle with a monotonic curve-domain luminance guidance that replaces multi-attribute pixel losses with one unified luminance energy in a low-dimensional curve space. Smoothness and shape regularizers act only on the curve parameters, so the sampler is driven by a single luminance-side energy rather than several heterogeneous losses defined on x_t. Specifically, LUCE comprises three lightweight components: (1) a unified luminance target that blends the luminance of the current clean prediction with an input-derived exposure prior using a time-dependent trade-off; (2) a monotonic luminance lookup table (LUT), regularized for smoothness and shape, that enforces the target and yields a global, interpretable brightness trajectory; and (3) a runtime chromatic decoupling transform using per-pixel scalar gains to realize the luminance edit while preserving hue. Experiments on paired and unpaired benchmarks validate these benefits.

In summary, we make three main contributions:

- We analyze limitations of existing training-free diffusion-based LLIE methods that rely on weighted pixel-space objectives, and identify gradient competition and luminance–chroma entanglement as core sources of instability and color shifts.
- We introduce **LUCE**, a training-free framework that constructs a unified luminance target and enforces it via a single monotonic curve-domain en-

ergy in a low-dimensional LUT parameterization, producing a smooth and interpretable brightness trajectory without multi-loss balancing.
– We propose a hue-preserving chromatic decoupling mechanism implemented as per-pixel scalar gains, improving chromatic fidelity during luminance editing.

2 Related Works

2.1 Low-Light Image Enhancement

Early low-light image enhancement (LLIE) methods rely on hand-crafted priors such as histogram equalization and Retinex-style illumination estimation. Representative approaches including LIME [7] estimate per-pixel illumination maps and refine them with structure priors to improve visibility. These methods are simple and efficient, but often amplify noise or introduce halos in complex real-world scenes.

Deep learning has substantially advanced LLIE by learning data-driven illumination priors. Supervised methods train convolutional or transformer-based networks on paired datasets such as LOL and LOLv2 [39, 45], either modeling illumination–reflectance decomposition [1, 25, 39, 40] or directly regressing enhanced images from RAW or RGB inputs [37]. While they achieve strong enhancement, they require labor-intensive data acquisition and may generalize poorly across cameras and capture conditions. To avoid paired data, zero-reference approaches such as Zero-DCE and its variants [6, 21] estimate image-specific tone curves from non-reference losses; later works refine curve design and loss balancing [23, 33]. These methods offer an interpretable curve formulation but still rely on offline training and can struggle under extremely dark or spatially heterogeneous low light. Many curve-based formulations operate with global or low-DOF mappings, which are interpretable but may be less flexible under strongly spatially varying illumination.

2.2 Diffusion Models for Image Enhancement and Restoration

Diffusion models have emerged as strong priors for low-level vision tasks such as denoising, super-resolution, and image-to-image translation [31, 32, 48]. In LLIE, several works train task-specific diffusion models to map low-light inputs to normally exposed images. DiffLL [14] adopts a wavelet-based conditional diffusion model for efficient enhancement, while LightenDiffusion [15] combines latent-space Retinex decomposition with an unsupervised diffusion backbone. Related frameworks incorporate Retinex priors, frequency-domain conditioning, or RAW processing (e.g., [9,13,19,46]) to improve perceptual quality and robustness. However, these approaches require training or fine-tuning diffusion models, which is costly and often tied to specific datasets or sensor domains.

A complementary line of research performs restoration directly on top of pre-trained diffusion backbones via training-free, gradient-based guidance at sam-

pling time. For instance, PGDiff [43] introduces partial guidance for face restoration by modeling structure and color statistics as auxiliary objectives, and AGLLDiff [24] extends this paradigm to real-world LLIE by designing pixel-space exposure, structure, and color attributes to steer sampling without paired data. While effective, these training-free methods typically aggregate multiple RGB-domain objectives with heuristic weights; their gradients can conflict and entangle luminance and chroma, which may cause unstable exposure control or hue shifts. We also note that a common practical baseline is to apply simple tone mapping (e.g., gamma) after denoising; our focus is on sampler-side guidance formulations that shape the denoising trajectory.

2.3 Training-Free Guidance on Pretrained Diffusion Backbones

Pretrained diffusion models can be adapted to editing/restoration without retraining by modifying the reverse process with auxiliary guidance, e.g., classifier and classifier-free guidance that inject objective gradients into sampling [4,11], or SDEdit-style stochastic denoising with noise reinitialization [28]. Along this line, training-free methods formulate restoration by adding energies on intermediate or output images during sampling.

For enhancement and color/tone manipulation, many works specify pixel-domain objectives and combine them with heuristic weights. PGDiff [43] aligns structure and color statistics, while AGLLDiff [24] steers LLIE with exposure/structure/color attributes in RGB space. Such designs are flexible but often suffer from weight sensitivity, luminance–chroma entanglement, and unstable exposure on challenging scenes. Other variants explore non-RGB cues, including illumination/Retinex priors [33,46], alternative color spaces [36,42], and frequency/RAW statistics [3,35,41], yet typically still aggregate multiple heterogeneous objectives whose gradients may interfere.

Within this training-free paradigm, LUCE differs in that it constrains luminance guidance to a low-dimensional monotone curve space and expresses exposure control as a single scalar energy in the curve domain, rather than combining multiple pixel-space objectives. The luminance edit is realized through explicit hue-preserving chromatic decoupling, separating luminance control from chromatic adjustment (details in Sec. 3).

3 Method

3.1 Preliminaries

We formulate training-free low-light enhancement as conditional sampling from a generative model $p(x_0 \mid x_{\text{in}})$, where x_{in} is a low-light input and x_0 denotes a well-exposed image. Our method builds on a pre-trained diffusion UNet and modifies only its inference procedure, while keeping all network parameters frozen.

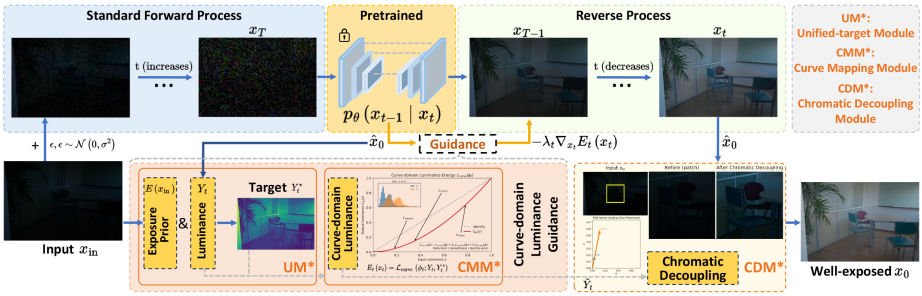

Fig. 2: Overview of LUCE. A frozen DDPM takes a low-light input x_{in} and produces a clean estimate $\hat{x}_0(x_t, t)$ whose luminance Y_t is fused with an exposure prior $E(x_{\text{in}})$ by the unified-target module (UM*) to form $Y_t^\star$. The curve mapping module (CMM*) fits a monotonic curve C_{ϕ_t} and defines a single curve-domain energy $E_t(x_t) = \mathcal{L}_{\text{curve}}(\phi_t; Y_t, Y_t^\star)$ whose gradient $-\lambda_t \nabla_{x_t} E_t(x_t)$ guides the discrete sampler, while the chromatic decoupling module (CDM*) realizes the edited luminance in RGB via a hue-preserving per-pixel scalar gain to obtain a well-exposed x_0.

Diffusion backbone. The UNet is trained in the standard DDPM setting. Given a clean image x_0, noisy states are generated as

$$x_t = \alpha_t x_0 + \sigma_t \varepsilon, \qquad \varepsilon \sim \mathcal{N}(0, I), \tag{1}$$

and the network $\varepsilon_\theta(x_t, t)$ predicts the added noise. A clean estimate is recovered by

$$\hat{x}_0(x_t, t) = \frac{x_t - \sigma_t\, \varepsilon_\theta(x_t, t)}{\alpha_t}. \tag{2}$$

We denote its luminance channel by

$$Y(\hat{x}_0) = 0.299\, \hat{x}_0^{(R)} + 0.587\, \hat{x}_0^{(G)} + 0.114\, \hat{x}_0^{(B)}. \tag{3}$$

In particular, luminance is always computed in linear RGB (i.e., we inverse-gamma convert from sRGB before computing $Y(\cdot)$), and results are gamma-encoded back to sRGB only for visualization.

Sampler-side guidance (discrete). At inference, we use a standard discrete DDPM sampler and inject an additional gradient-based correction derived from a single scalar energy $E_t(x_t)$ (defined later). Denoting the vanilla DDPM update by $\text{DDPMUpdate}_\theta(\cdot)$, we apply

$$x_{t-1} \;=\; \text{DDPMUpdate}_\theta(x_t, t) \;-\; \lambda_t\, \nabla_{x_t} E_t(x_t), \tag{4}$$

with a time-dependent guidance strength $\lambda_t \geq 0$. The backbone parameters and the base update rule remain unchanged, and we typically apply the correction to the last K steps (we set K=10 in all experiments; see Sec. 4.2).

3.2 LUCE Pipeline

Fig. 2 overviews LUCE. Given a low-light input x_{in}, we run a pre-trained diffusion UNet with a standard DDPM [10] sampler. Rather than attaching multiple pixel-space objectives to x_t, we cast sampler-side guidance as constrained curve-domain luminance control: luminance cues are consolidated into a single scalar energy defined in a low-dimensional monotone curve space, whose gradient is injected into the discrete sampler via Eq. (4), while keeping both the backbone and the DDPM rule frozen.

LUCE defines its guidance objective in the luminance domain. First, it constructs a unified luminance target $Y_t^\star$ by fusing the luminance of the current clean estimate with an exposure prior derived from x_{in} (Sec. 3.2). Second, it fits a monotonic luminance curve in the 1D luminance domain so that the curve output follows $Y_t^\star$, with simple priors to keep the curve smooth and close to identity (Sec. 3.2). Third, a runtime chromatic decoupling module applies a per-pixel hue-preserving color transform that realizes the desired luminance changes in RGB space (Sec. 3.2). The unified target and curve mapping define a single curve-domain energy, while chromatic decoupling realizes the resulting luminance edit in RGB space.

Unified Luminance Target. Training-free LLIE methods on diffusion backbones typically impose several luminance-related pixel-space objectives (e.g., exposure alignment, preservation of well-exposed regions, global brightness statistics). Optimized in parallel, their gradients can point in conflicting directions and require careful weight tuning. LUCE instead collapses all luminance cues into a single unified luminance target that serves as the brightness reference along the sampling trajectory.

Let x_t be the current noisy state and $\hat{x}_0(x_t, t)$ the corresponding clean estimate from the diffusion UNet. Its luminance map is

$$Y_t = Y\big(\hat{x}_0(x_t, t)\big) \in \mathbb{R}^{H\times W}, \tag{5}$$

computed via Eq. (3). In parallel, we derive an exposure prior $E(x_{\text{in}}) \in \mathbb{R}^{H\times W}$ from the input. In our default instantiation, we compute input luminance $Y_{\text{in}} = Y(x_{\text{in}})$ in linear RGB, apply a global gamma correction $\hat{Y}_\gamma(p) = Y_{\text{in}}(p)^\gamma$ with $\gamma \in (0, 1)$, and then perform a per-image affine rescaling with clipping:

$$E(x_{\text{in}})(p) = \text{clamp}_{[0,1]}\big(a\,\hat{Y}_\gamma(p) + b\big),$$

where (a, b) are chosen to match a fixed target mean luminance μ_{tar}. We use a single configuration across datasets, with γ=0.6 and μ_{tar}=0.5; implementation details are provided in the supplementary.

We define the unified luminance target at time t as a convex combination

$$Y_t^\star = \beta_t\, Y_t + (1 - \beta_t)\, E(x_{\text{in}}), \qquad \beta_t \in [0, 1], \tag{6}$$

applied element-wise. The schedule β_t trades off backbone fidelity and exposure prior: earlier guided steps use smaller β_t to brighten dark regions; as $t \to 0$,

β_t increases to preserve generative details. We use a fixed monotone schedule where β_t increases toward 1 as $t \to 0$ (e.g., up to 0.9 in our default setting), so the exposure prior mainly affects early steps while late steps are increasingly anchored by the diffusion prediction.

Curve-Guided Luminance Energy. A pixel-wise penalty $\|Y_t - Y_t^\star\|$ tends to apply independent local corrections and may distort global exposure/contrast. LUCE instead uses a global 1D luminance mapping, represented as a monotonic curve, so that luminance edits are coordinated by a single, interpretable function.

Monotonic luminance curve. Let $y \in [0,1]$ denote a normalized luminance value from Y_t. We define a parametric curve

$$C_\phi : [0,1] \to [0,1], \tag{7}$$

implemented as a piecewise-linear lookup table with uniform knots $\{u_b\}_{b=0}^{B}$ and values $\{\phi_b\}_{b=0}^{B}$. For any $y \in [u_b, u_{b+1}]$,

$$C_\phi(y) = \frac{u_{b+1} - y}{u_{b+1} - u_b}\,\phi_b + \frac{y - u_b}{u_{b+1} - u_b}\,\phi_{b+1}. \tag{8}$$

To preserve luminance ordering and avoid structure inversion, we impose the monotonicity constraint

$$0 = \phi_0 \le \phi_1 \le \cdots \le \phi_B = 1, \tag{9}$$

and obtain the curve-edited luminance $\tilde{Y}_t = C_\phi(Y_t)$.

Energy in the curve domain. We align the curve output with the unified target via

$$\mathcal{L}_{\text{match}}(\phi) = \frac{1}{N}\sum_{p=1}^{N} \big|C_\phi\big(Y_t(p)\big) - Y_t^\star(p)\big|, \tag{10}$$

and stabilize the mapping with two lightweight regularizers: an ℓ_1 penalty on second-order differences for smoothness and an ℓ_2 penalty toward the identity mapping $y \mapsto y$ (identity regularization). Denote them by $\mathcal{L}_{\text{smooth}}(\phi)$ and $\mathcal{L}_{\text{shape}}(\phi)$. The single curve-domain energy is

$$\mathcal{L}_{\text{curve}}(\phi) = \mathcal{L}_{\text{match}}(\phi) + \lambda_1\,\mathcal{L}_{\text{smooth}}(\phi) + \lambda_2\,\mathcal{L}_{\text{shape}}(\phi), \tag{11}$$

where $\lambda_1, \lambda_2 \ge 0$ are fixed. We optimize ϕ in the low-dimensional curve space ($B \ll N$) with a few projected gradient steps on ϕ (e.g., isotonic projection) that enforce (9); this is applied only in the last K guided steps (runtime is reported in Sec. 4.2). During sampling, ϕ_t is treated as fixed when differentiating w.r.t. x_t, so $\nabla_{x_t} E_t(x_t)$ is driven solely by the luminance-matching term $\mathcal{L}_{\text{match}}$.

Scalar guidance from the curve. Given ϕ_t, we define a scalar guidance on the current state x_t as

$$E_t(x_t) = \mathcal{L}_{\text{curve}}\big(\phi_t; Y_t, Y_t^\star\big), \tag{12}$$

where $Y_t = Y\big(\hat{x}_0(x_t, t)\big)$ depends on x_t through the clean estimate. In inference, ϕ_t is treated as fixed when differentiating $E_t(x_t)$ w.r.t. x_t. The sampler update then follows Eq. (4), injecting a single guidance term defined in the luminance domain instead of a weighted sum of heterogeneous pixel-space gradients. In other words, when taking $\nabla_{x_t} E_t(x_t)$ the curve regularizers no longer contribute, and the sampler is driven by one luminance-only energy rather than multiple objectives defined directly on x_t.

Runtime Chromatic Decoupling. The curve-domain energy specifies how luminance should evolve, but not how RGB channels should be updated to realize this change. Naively adjusting channels independently can cause hue shifts and over-saturation, especially in dark regions. LUCE therefore applies luminance edits through a simple per-pixel linear transform that decouples brightness from chroma.

Let $c_t(p) \in [0,1]^3$ be the RGB vector of $\hat{x}_0(x_t, t)$ at pixel p, and

$$Y_t(p) = w^\top c_t(p), \quad w = (0.299,\, 0.587,\, 0.114)^\top. \tag{13}$$

After optimizing ϕ_t, the desired luminance is

$$\tilde{Y}_t(p) = C_{\phi_t}\big(Y_t(p)\big). \tag{14}$$

We realize this change with a per-pixel scalar gain:

$$\tilde{c}_t(p) = s_t(p)\, c_t(p), \qquad s_t(p) = \frac{\tilde{Y}_t(p)}{\max\big(Y_t(p), \varepsilon\big)}, \tag{15}$$

with a small $\varepsilon > 0$. When $Y_t(p) > \varepsilon$ and the gain is not clipped, the transform exactly realizes $w^\top \tilde{c}_t(p) = \tilde{Y}_t(p)$. Otherwise, it provides a bounded approximation. Any positive scalar gain preserves the RGB direction and hence the hue before gamut clipping. This chromatic decoupling realizes the curve-edited luminance in RGB space while suppressing color casts and over-saturation.

Curve-Domain Luminance Guidance in the Sampler. The components above define the complete LUCE inference pipeline on a frozen diffusion backbone.

At time t, the diffusion UNet produces $\hat{x}_0(x_t, t)$, and we compute $Y_t = Y\big(\hat{x}_0(x_t, t)\big)$. The unified target $Y_t^\star$ follows Eq. (6); the curve parameters ϕ_t are obtained by minimizing $\mathcal{L}_{\text{curve}}(\phi; Y_t, Y_t^\star)$ under (9). The scalar energy $E_t(x_t)$ in Eq. (12) forms the guidance term $-\lambda_t \nabla_{x_t} E_t(x_t)$ in Eq. (4). The chromatic transform in Sec. 3.2 realizes the curve-edited luminance in RGB space while preserving hue. The backbone parameters and the base DDPM update remain unchanged.

Conceptually, LUCE reconciles different luminance desiderata (backbone-consistent prediction, exposure prior, and curve simplicity) before entering the sampler by projecting them onto a unified target and a single monotonic curve. Consequently, $\nabla_{x_t} E_t(x_t)$ represents a coherent, luminance-guided adjustment on top of the pretrained diffusion prior, rather than a weighted sum of competing pixel-space gradients, yielding stable, interpretable exposure control.

In summary, LUCE injects a luminance-driven, curve-domain single-energy guidance into a frozen diffusion backbone via a discrete, gradient-based correction. All luminance cues are first reconciled into a unified target, realized by a monotonic curve with lightweight regularization, and mapped back to RGB through a hue-preserving transform. This reduces objective interference by constraining guidance to a low-dimensional monotone curve space while remaining fully training-free.

4 Experiments

4.1 Datasets and Evaluation Metrics

Paired low-light datasets. We evaluate LUCE on three representative paired LLIE benchmarks: LOL-v1 [39], LOL-v2 (Real and Synthetic) [45], and LSRW [8]. LOL-v1 contains 500 low-/normal-light pairs (485/15 for train/test). LOL-v2-Real and LOL-v2-Synthetic provide real-captured and synthetically darkened pairs with diverse indoor and outdoor scenes, and LSRW offers real-world pairs with mixed degradations such as noise and color cast; for all three, we follow the standard train/test splits used in prior work. On these paired datasets, we report PSNR and SSIM [38] for reconstruction fidelity and LPIPS [49] for perceptual similarity (higher PSNR/SSIM and lower LPIPS are better). For LOL-v2, we report the combined Real+Synthetic score in the main paper and the two subsets separately in the supplementary.

Unpaired low-light datasets. To assess generalization without ground-truth references, we further evaluate on five standard unpaired benchmarks: LIME [7], NPE [35], MEF [26], DICM [20], and VV [34], which cover diverse real-world low-light scenes but provide only input images. Following prior work, we use BRISQUE [29] and noise level (NL) [2] as no-reference indicators of perceptual quality and residual noise, where lower scores indicate better performance.

4.2 Implementation Details

LUCE is built on a 256×256 DDPM denoising UNet pretrained on ImageNet, following the public 1000-step noise schedule. At inference, the backbone is frozen and we only modify the last 10 reverse steps by adding our luminance energy to the sampler's gradient-based guidance; earlier steps follow the original sampler. All images are resized to 256×256 for sampling and bicubically rescaled back to their original resolution for evaluation.

Table 1: Quantitative comparison on LOL-v1, LOL-v2 (Real+Synthetic combined), and LSRW. Methods are grouped by supervision type (**S**: supervised, **U**: unsupervised). We report PSNR/SSIM/LPIPS (P/S/L). Best and second-best results within the **U** group are highlighted in boldface and with underlining, respectively.

Type	Method	LOL-v1 (P/S/L)	LOL-v2 (R+S) (P/S/L)	LSRW (P/S/L)
S	Retinex-Net [39]	17.62/0.64/0.38	16.62/0.58/0.40	15.58/0.41/0.39
	KinD [51]	21.02/0.84/0.20	17.93/0.73/0.31	15.51/0.38/0.39
	KinD++ [50]	17.75/0.76/0.20	18.05/0.73/0.30	16.09/0.39/0.37
	URetinex-Net [40]	19.84/0.82/0.13	20.56/0.86/0.24	18.27/0.52/0.30
	Retinexformer [1]	25.15/0.84/0.13	20.23/0.80/0.25	19.23/0.54/0.31
	DiffLL [14]	27.93/0.88/0.11	25.47/0.89/0.11	19.27/0.55/0.30
	GSAD [12]	22.02/0.85/0.14	19.31/0.89/0.13	17.41/0.51/0.29
U	EnlightenGAN [16]	17.48/0.65/0.32	15.65/0.58/0.49	17.05/0.46/0.33
	Zero-DCE [6]	14.86/0.55/0.34	16.52/0.68/0.26	15.84/0.45/**0.31**
	Zero-DCE++ [21]	15.32/0.56/0.33	17.11/0.70/0.28	15.32/0.49/0.33
	ExCNet [47]	16.05/0.63/0.37	14.40/0.57/0.46	14.26/0.46/0.50
	RUAS [25]	16.44/0.49/0.27	15.32/0.58/0.49	14.31/0.48/0.47
	SCI [27]	14.78/0.53/0.33	13.38/0.61/0.33	15.24/0.42/0.45
	NeRCo [44]	19.81/0.74/0.24	17.19/0.64/0.39	**18.82**/0.51/0.34
	CLIP-LIT [22]	12.39/0.49/0.38	15.20/0.55/0.58	13.46/0.40/0.35
	PairLIE [5]	19.46/0.74/0.25	17.75/0.75/0.31	17.59/0.49/0.33
	CoLIE [3]	20.28/0.49/0.37	15.39/0.59/0.29	16.21/0.48/0.34
	LightenDiffusion [15]	20.44/0.81/0.19	19.24/<u>0.83</u>/**0.17**	18.54/0.54/**0.31**
	QuadPrior [36]	**22.85**/0.80/0.20	18.62/0.81/0.21	17.99/<u>0.55</u>/0.37
	AGLLDiff [24]	21.81/<u>0.84</u>/<u>0.16</u>	<u>19.59</u>/<u>0.83</u>/0.21	16.92/**0.56**/<u>0.32</u>
	LUCE (Ours)	<u>22.81</u>/**0.85**/**0.15**	**20.21**/**0.84**/<u>0.19</u>	<u>18.80</u>/**0.56**/<u>0.32</u>

The luminance curve is implemented as a piecewise-linear lookup table with B=32 bins. Unless otherwise stated, we fix the curve regularization weights to λ_1=0.01 (smoothness) and λ_2=0.1 (shape), use a monotonically increasing schedule for β_t (from 0.3 to 0.9, increasing toward 1 as $t \to 0$) and guidance strength λ_t (from 0 to 1.0), and apply chromatic decoupling with ε=10^{-3} and per-pixel gain $s_t(p)$ clipped to $[0.5, 3.0]$. The same hyperparameters are used for all datasets. All experiments are training-free and run on a single NVIDIA A100 GPU. Reported runtimes include curve fitting and guidance computations.

4.3 Benchmarking Results

Quantitative Analysis on Low-Light Enhancement. Table 1 compares LUCE with supervised and unsupervised LLIE methods on three paired benchmarks. Inference times reported in Tab. 2 are measured per 256×256 image on a single NVIDIA A100 GPU. Among unsupervised approaches, LUCE consistently achieves top-tier results (best or second-best) across the three paired benchmarks. On LOL-v1, it attains the best SSIM/LPIPS and second-best PSNR, indicating that our monotonic curve-domain guidance improves structural and perceptual fidelity while remaining competitive in distortion. On LOL-v2 (Real+Synthetic combined), LUCE achieves the best PSNR/SSIM and the second-best LPIPS, demonstrating a favorable distortion–perception trade-off relative to AGLLDiff and LightenDiffusion; this suggests that consolidating luminance guidance into a single curve-domain energy yields more stable exposure control

Table 2: Efficiency-aware comparison on unpaired datasets (LIME, NPE, MEF, DICM, VV). Efficiency metrics include model parameters (M), MACs (G), and inference time (s). BRISQUE and noise level (NL) evaluate perceptual quality and residual noise (lower is better). Best and second-best results within the **U** group are highlighted in boldface and with underlining, respectively.

Type	Method	Efficiency			No-reference Quality	
		Params(M)	MACs(G)	Time(s)	BRISQUE↓	NL↓
S	Retinex-Net [39]	0.555	87.28	0.093	27.10	3.25
	KinD [51]	8.02	159.56	0.096	26.89	0.70
	KinD++ [50]	8.27	258.62	0.694	26.16	0.43
	URetinex-Net [40]	0.341	229.0	0.248	23.80	1.32
	Retinexformer [1]	1.61	68.39	0.310	14.77	1.06
	DiffLL [14]	22.08	21.88	0.340	24.13	0.68
U	EnlightenGAN [16]	8.64	65.88	0.030	20.65	0.79
	Zero-DCE [6]	0.079	20.92	0.016	21.76	1.57
	Zero-DCE++ [21]	0.011	0.021	0.003	19.34	1.15
	ExCNet [47]	8.27	-	14.62	19.03	1.56
	RUAS [25]	0.003	0.951	0.019	29.91	2.09
	SCI [27]	0.00026	0.095	0.002	**16.73**	0.85
	NeRCo [44]	23.30	822.8	0.949	22.81	**0.65**
	CLIP-LIT [22]	0.279	73.04	0.054	23.44	1.96
	PairLIE [5]	0.342	89.99	0.056	29.84	1.47
	LightenDiffusion [15]	26.54	2257.42	9.539	18.53	1.11
	QuadPrior [36]	1252.75	-	9.961	19.96	1.34
	AGLLDiff [24]	552.81	1114.68	20.43	19.55	0.80
	LUCE (Ours)	556.22	1106.35	10.53	<u>18.03</u>	<u>0.76</u>

and improved color fidelity compared with multi-objective pixel-space guidance. On LSRW, LUCE slightly trails NeRCo in PSNR but achieves the best SSIM (tied) and near-best LPIPS, suggesting that our more conservative luminance editing avoids over-brightening that can inflate PSNR at the cost of perceptual quality. Compared with AGLLDiff, LUCE improves PSNR on all three datasets and achieves equal or better SSIM/LPIPS, with almost identical parameter and MAC counts but roughly half the inference time. On unpaired benchmarks, LUCE further attains the second-best BRISQUE and the second-best NL within the unsupervised group (Tab. 2), suggesting improved perceptual quality with controlled residual noise. While lightweight curve-based CNNs remain faster by orders of magnitude, they exhibit noticeable performance gaps, highlighting that our training-free diffusion framework strikes a favorable balance between quality and computational cost.

Qualitative Comparison. As shown in Fig. 3, LUCE produces visually pleasing results that are closer to the ground truth than both training-based curves and existing training-free diffusion baselines. Zero-DCE and SCI brighten the scene but either wash out details or introduce noticeable color casts, while CLIP-LIT and CoLIE tend to generate greenish tones and uneven exposure across the image. The training-free AGLLDiff yields competitive brightness, yet the red-boxed regions reveal residual color bias and locally over-enhanced areas (e.g., around the target board and color wheel). In contrast, LUCE simultaneously restores

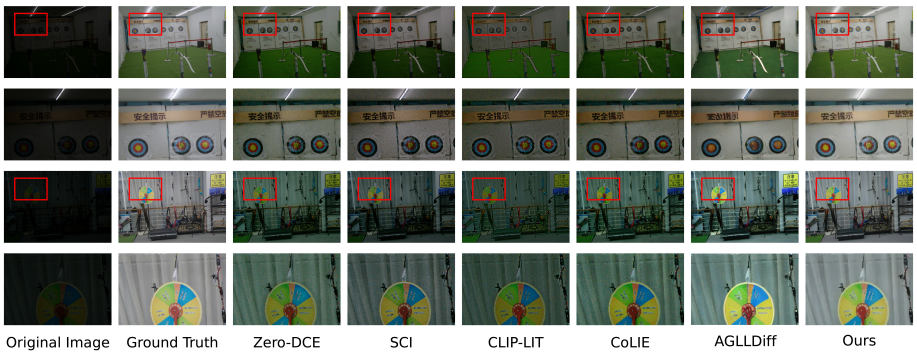

Fig. 3: Visual comparison on low-light images. Compared with representative curve-based and training-free diffusion baselines, LUCE produces better-exposed and less noisy results while preserving color fidelity and fine structures in challenging dark regions.

Table 3: Ablation study of LUCE components. Paired metrics are evaluated on LOL-v2, while no-reference metrics are averaged over unpaired datasets (LIME, NPE, MEF, DICM, VV). ↑ / ↓ indicate higher/lower is better.

Variant	Paired (LOL-v2)			No-reference	
	PSNR↑	SSIM↑	LPIPS↓	BRISQUE↓	NL↓
w/o UM*	19.70	0.832	0.205	20.71	1.45
w/o CMM*	19.85	0.836	0.200	19.86	1.39
w/o CDM*	19.78	0.834	0.208	19.03	1.08
LUCE (full)	**20.21**	**0.840**	**0.192**	**18.03**	**0.76**

global brightness and maintains consistent hue: text and edges inside the red boxes remain sharp, the background illumination is smooth, and saturated colors (targets, wheel) are recovered without overshooting. These observations corroborate our design: the single curve-domain luminance energy enforces a stable monotonic exposure trajectory, and the chromatic decoupling module prevents the hue distortions commonly seen in multi-attribute RGB-space guidance.

4.4 Ablation Studies

Core components on paired data. Table 3 compares three key variants of LUCE on LOL-v2. Removing the unified luminance target (w/o UM*) causes the largest PSNR and SSIM drops, showing the importance of an explicit brightness reference. Removing the curve mapping module (w/o CMM*) leads to a moderate degradation, indicating that explicit monotonic curve fitting is important for enforcing globally consistent exposure transitions rather than unstable local edits. Discarding chromatic decoupling (w/o CDM*) causes the largest LPIPS degradation, confirming its importance for perceptual and chromatic fidelity. The full

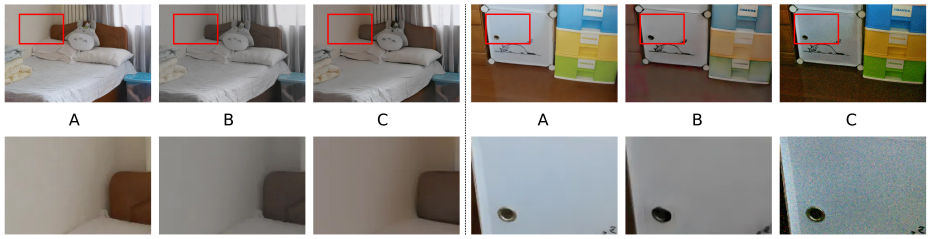

Fig. 4: Visual ablation of LUCE. For each scene, column A shows the full LUCE, column B removes the curve mapping module (w/o CMM*, replacing curve-domain guidance with a pixel-wise ℓ_1 loss on luminance), and column C removes chromatic decoupling (w/o CDM*, without the hue-preserving scalar-gain transform). Zoom-in regions highlight that B produces flat, washed-out surfaces, while C introduces color casts and amplified noise, whereas A preserves both natural brightness and clean, faithful colors.

LUCE, which combines UM*, CMM*, and CDM*, achieves the best PSNR/SSIM and the lowest LPIPS, validating the effectiveness of the proposed design.

Effect on unpaired benchmarks. On the unpaired datasets (LIME, NPE, MEF, DICM and VV), Table 3 shows that the full LUCE achieves the best BRISQUE/NL among its ablated variants, indicating better perceptual quality and fewer residual artifacts when UM*, CMM*, and CDM* are all enabled. Removing UM* yields clearly higher BRISQUE/NL, reflecting unstable exposure and amplified noise when the guidance lacks an explicit brightness reference. Removing CMM* also deteriorates BRISQUE, suggesting that monotonic curve fitting helps avoid over-stretched contrast and unnatural tone curves. Finally, omitting CDM* increases NL and worsens BRISQUE, matching visual observations that removing hue-preserving chromatic decoupling produces noisier and less color-consistent outputs. In the method-level comparison in Table 2, LUCE achieves the second-best BRISQUE and the second-best NL among unsupervised methods, indicating a favorable quality–noise trade-off under a training-free setting.

Ablation visualization. Fig. 4 provides a visual ablation of the key modules. Column A is the full LUCE, column B is w/o CMM* (replacing curve-domain guidance with a pixel-wise ℓ_1 loss on luminance), and column C is w/o CDM* (without the hue-preserving scalar-gain transform). Compared with A, variant B yields flatter walls and cabinet surfaces with reduced local contrast, showing that the global monotonic curve is crucial for coordinated exposure control. Variant C reaches similar brightness but suffers from visible color casts and strong chroma noise in dark regions, confirming the role of CDM* in hue-preserving luminance editing. We do not show w/o UM* here, as its main effect is a global exposure drift that is already reflected by the quantitative gaps in Table 3.

5 Conclusion

We presented LUCE, a training-free low-light enhancement framework that performs constrained curve-domain luminance guidance with hue-preserving chromatic decoupling on a frozen diffusion backbone, enabling stable and controllable exposure adjustment with improved chromatic fidelity over prior training-free diffusion baselines. Beyond quantitative gains, LUCE provides an interpretable luminance control interface during sampling via a low-dimensional monotone curve, which helps stabilize exposure evolution compared with weighted multi-objective pixel-space guidance. Despite these advantages, the current framework still relies on hand-crafted exposure priors and fixed guidance schedules. Making these components more adaptive and extending LUCE to higher-resolution diffusion backbones, as well as RAW and video settings, are promising directions for future work.

Acknowledgements

This work was supported by the Beijing Smart Agriculture Innovation Consortium Project (BAIC10-2026).

References

1. Cai, Y., Bian, H., Lin, J., Wang, H., Timofte, R., Zhang, Y.: Retinexformer: One-stage retinex-based transformer for low-light image enhancement. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 12504–12513 (2023)
2. Chen, G., Zhu, F., Heng, P.A.: An efficient statistical method for image noise level estimation. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 477–485 (2015)
3. Chobola, T., Liu, Y., Zhang, H., Schnabel, J.A., Peng, T.: Fast context-based low-light image enhancement via neural implicit representations. In: European Conference on Computer Vision. pp. 413–430 (2024)
4. Dhariwal, P., Nichol, A.: Diffusion models beat gans on image synthesis. Advances in Neural Information Processing Systems **34**, 8780–8794 (2021)
5. Fu, Z., Yang, Y., Tu, X., Huang, Y., Ding, X., Ma, K.K.: Learning a simple low-light image enhancer from paired low-light instances. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 22252–22261 (2023)
6. Guo, C., Li, C., Guo, J., Loy, C.C., Hou, J., Kwong, S., Cong, R.: 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)
7. Guo, X., Li, Y., Ling, H.: Lime: Low-light image enhancement via illumination map estimation. IEEE Transactions on Image Processing **26**(2), 982–993 (2017)
8. Hai, J., Xuan, Z., Yang, R., Hao, Y., Zou, F., Lin, F., Han, S.: R2rnet: Low-light image enhancement via real-low to real-normal network. Journal of Visual Communication and Image Representation **90**, 103712 (2023)

9. He, J., Xue, M., Ning, A., Song, C.: Zero-reference lighting estimation diffusion model for low-light image enhancement. arXiv preprint arXiv:2403.02879 (2024)
10. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. Advances in Neural Information Processing Systems **33**, 6840–6851 (2020)
11. Ho, J., Salimans, T.: Classifier-free diffusion guidance. arXiv preprint arXiv:2207.12598 (2022)
12. Hou, J., Zhu, Z., Hou, J., Liu, H., Zeng, H., Yuan, H.: Global structure-aware diffusion process for low-light image enhancement. Advances in Neural Information Processing Systems **36**, 79734–79747 (2023)
13. Huang, Y., Liao, X., Liang, J., Shi, B., Xu, Y., Le Callet, P.: Detail-preserving diffusion models for low-light image enhancement. IEEE Transactions on Circuits and Systems for Video Technology **35**(4), 3396–3409 (2025)
14. Jiang, H., Luo, A., Fan, H., Han, S., Liu, S.: Low-light image enhancement with wavelet-based diffusion models. ACM Transactions on Graphics **42**(6), 1–14 (2023)
15. Jiang, H., Luo, A., Liu, X., Han, S., Liu, S.: Lightendiffusion: Unsupervised low-light image enhancement with latent-retinex diffusion models. In: European Conference on Computer Vision. pp. 161–179 (2024)
16. Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., Yang, J., Zhou, P., Wang, Z.: Enlightengan: Deep light enhancement without paired supervision. IEEE Transactions on Image Processing **30**, 2340–2349 (2021)
17. Kim, W.: Low-light image enhancement: A comparative review and prospects. IEEE Access **10**, 84535–84557 (2022)
18. Kong, T., Rey, W.P.: Self-supervised low-light image enhancement based on retinex and histogram equalization prior. In: Proceedings of the International Conference on Image Processing, Machine Learning and Pattern Recognition. pp. 143–147 (2024)
19. Lan, G., Ma, Q., Yang, Y., Wang, Z., Wang, D., Li, X., Zhao, B.: Efficient diffusion as low light enhancer. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 21277–21286 (2025)
20. Lee, C., Lee, C., Kim, C.S.: Contrast enhancement based on layered difference representation of 2d histograms. IEEE Transactions on Image Processing **22**(12), 5372–5384 (2013)
21. Li, C., Guo, C., Loy, C.C.: Learning to enhance low-light image via zero-reference deep curve estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(8), 4225–4238 (2022)
22. Liang, Z., Li, C., Zhou, S., Feng, R., Loy, C.C.: Iterative prompt learning for unsupervised backlit image enhancement. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 8094–8103 (2023)
23. Lin, Y., Fu, Z., Wen, K., Ye, T., Chen, S., Meng, G., Wang, Y., Kong, C., Huang, Y., Tu, X., Ding, X.: Dplut: Unsupervised low-light image enhancement with lookup tables and diffusion priors. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 39, pp. 5316–5324 (2025)
24. Lin, Y., Ye, T., Chen, S., Fu, Z., Wang, Y., Chai, W., Xing, Z., Li, W., Zhu, L., Ding, X.: Aglldiff: Guiding diffusion models towards unsupervised training-free real-world low-light image enhancement. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 39, pp. 5307–5315 (2025)
25. Liu, R., Ma, L., Zhang, J., Fan, X., Luo, Z.: Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 10561–10570 (2021)

26. Ma, K., Zeng, K., Wang, Z.: Perceptual quality assessment for multi-exposure image fusion. IEEE Transactions on Image Processing **24**(11), 3345–3356 (2015)
27. Ma, L., Ma, T., Liu, R., Fan, X., Luo, Z.: Toward fast, flexible, and robust low-light image enhancement. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 5637–5646 (2022)
28. Meng, C., He, Y., Song, Y., Song, J., Wu, J., Zhu, J.Y., Ermon, S.: Sdedit: Guided image synthesis and editing with stochastic differential equations. arXiv preprint arXiv:2108.01073 (2021)
29. Mittal, A., Moorthy, A.K., Bovik, A.C.: No-reference image quality assessment in the spatial domain. IEEE Transactions on Image Processing **21**(12), 4695–4708 (2012)
30. Qu, J., Liu, R.W., Gao, Y., Guo, Y., Zhu, F., Wang, F.Y.: Double domain guided real-time low-light image enhancement for ultra-high-definition transportation surveillance. IEEE Transactions on Intelligent Transportation Systems **25**(8), 9550–9562 (2024)
31. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 10684–10695 (2022)
32. Saharia, C., Chan, W., Chang, H., Lee, C., Ho, J., Salimans, T., Fleet, D., Norouzi, M.: Palette: Image-to-image diffusion models. In: ACM SIGGRAPH 2022 Conference Proceedings. pp. 1–10 (2022)
33. Shi, Y., Liu, D., Zhang, L., Tian, Y., Xia, X., Fu, X.: Zero-ig: Zero-shot illumination-guided joint denoising and adaptive enhancement for low-light images. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3015–3024 (2024)
34. Vonikakis, V., Kouskouridas, R., Gasteratos, A.: On the evaluation of illumination compensation algorithms. Multimedia Tools and Applications **77**, 9211–9231 (2018)
35. Wang, S., Zheng, J., Hu, H.M., Li, B.: Naturalness preserved enhancement algorithm for non-uniform illumination images. IEEE Transactions on Image Processing **22**(9), 3538–3548 (2013)
36. Wang, W., Yang, H., Fu, J., Liu, J.: Zero-reference low-light enhancement via physical quadruple priors. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 26057–26066 (2024)
37. Wang, Y., Wan, R., Yang, W., Li, H., Chau, L.P., Kot, A.: Low-light image enhancement with normalizing flow. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 36, pp. 2604–2612 (2022)
38. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing **13**(4), 600–612 (2004)
39. Wei, C., Wang, W., Yang, W., Liu, J.: Deep retinex decomposition for low-light enhancement. arXiv preprint arXiv:1808.04560 (2018)
40. Wu, W., Weng, J., Zhang, P., Wang, X., Yang, W., Jiang, J.: Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 5901–5910 (2022)
41. Xu, X., Wang, R., Fu, C.W., Jia, J.: Snr-aware low-light image enhancement. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 17714–17724 (2022)

42. Yan, Q., Feng, Y., Zhang, C., Pang, G., Shi, K., Wu, P., Dong, W., Sun, J., Zhang, Y.: Hvi: A new color space for low-light image enhancement. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 5678–5687 (2025)
43. Yang, P., Zhou, S., Tao, Q., Loy, C.C.: Pgdiff: Guiding diffusion models for versatile face restoration via partial guidance. Advances in Neural Information Processing Systems **36**, 32194–32214 (2023)
44. Yang, S., Ding, M., Wu, Y., Li, Z., Zhang, J.: Implicit neural representation for cooperative low-light image enhancement. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 12918–12927 (2023)
45. Yang, W., Wang, W., Huang, H., Wang, S., Liu, J.: Sparse gradient regularized deep retinex network for robust low-light image enhancement. IEEE Transactions on Image Processing **30**, 2072–2086 (2021)
46. Yi, X., Xu, H., Zhang, H., Tang, L., Ma, J.: Diff-retinex: Rethinking low-light image enhancement with a generative diffusion model. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 12302–12311 (2023)
47. Zhang, L., Zhang, L., Liu, X., Shen, Y., Zhang, S., Zhao, S.: Zero-shot restoration of back-lit images using deep internal learning. In: Proceedings of the 27th ACM International Conference on Multimedia. pp. 1623–1631 (2019)
48. Zhang, L., Rao, A., Agrawala, M.: Adding conditional control to text-to-image diffusion models. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 3836–3847 (2023)
49. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 586–595 (2018)
50. Zhang, Y., Guo, X., Ma, J., Liu, W., Zhang, J.: Beyond brightening low-light images. International Journal of Computer Vision **129**(4), 1013–1037 (2021)
51. Zhang, Y., Zhang, J., Guo, X.: Kindling the darkness: A practical low-light image enhancer. In: Proceedings of the 27th ACM International Conference on Multimedia. pp. 1632–1640 (2019)