InterLight: Leveraging Intrinsic Illumination Priors for Low-Light Image Enhancement

Ziqi Wang[1], Xu Zhang[1]*, Laibin Chang[1], Shi Chen[2], Jiaqi Ma[3], and Huan Zhang[4]†
[1]National Engineering Research Center for Multimedia Software, School of Computer Science, Wuhan University
[2]Department of Computer Science, University of Macau
[3]Department of Computer Vision, Mohamed bin Zayed University of Artificial Intelligence
[4]School of Information Engineering, Guangdong University of Technology
{2023302112030, zhangx0802, changlb666}@whu.edu.cn, chenshi@um.edu.mo,
jiaqi.ma@mbzuai.ac.ae, huanzhang2021@gdut.edu.cn

Abstract

Low-Light Image Enhancement (LLIE) has long been a challenging problem in low-level vision, as insufficient illumination often leads to low contrast, detail loss, and noise. Recent studies show that deep learning-based Retinex theory can effectively decouple illumination and reflectance. However, existing methods frequently suffer from over-enhancement or color distortion, and often assume uniform noise or ideal lighting. To address these limitations, we propose InterLight, a novel framework that systematically excavates and operationalizes intrinsic illumination priors for LLIE. Our core insight is that robust enhancement requires not just estimating illumination, but constructing an illumination-aware pipeline. We first inject sensor-level illumination-response priors via physics-guided augmentation, then represent the degradation through adaptive prompts conditioned on the scene's latent illumination state. This explicit representation directly guides a luminance-gated intrinsic memory mechanism to selectively compensate for information loss, prioritizing reconstruction in dark regions while preserving fidelity in bright ones. Finally, the entire process is regularized by a self-supervised consistency objective that distills illumination-invariant features. By deeply exploiting intrinsic illumination priors, our method achieves clearer textures and more visually coherent enhancement results. Extensive experiments across multiple benchmarks demonstrate the effectiveness of our approach. Code is available at: https://github.com/House-yuyu/InterLight.

1 Introduction

Low-Light Image Enhancement (LLIE) [Kimmel *et al.*, 2003; Arici *et al.*, 2009; Guo *et al.*, 2016; Wei *et al.*, 2018; Li *et al.*, 2018; Guo *et al.*, 2020; Ma *et al.*, 2022; Cai *et al.*, 2023; Ma *et al.*, 2025; Zhang *et al.*, 2025a] is a fundamental challenge in computer vision, as images captured under insufficient illumination often suffer from severe noise, low contrast, color distortion, and structural detail loss. These degradations not only compromise visual quality but also hinder the performance of downstream tasks such as detection [Wang *et al.*, 2025], segmentation [Gu *et al.*, 2025], and autonomous driving [Wu *et al.*, 2025].

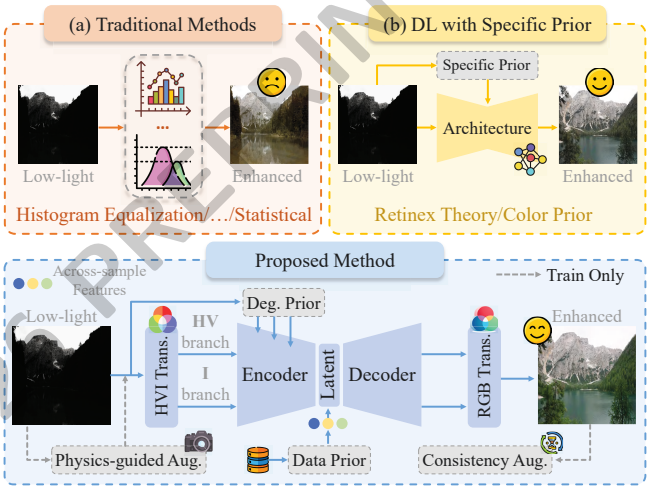

Figure 1: Comparison of InterLight with traditional and deep learning–based with specific prior LLIE methods. (a) Traditional approaches using handcrafted priors (e.g., histogram equalization, Retinex) often fail under non-uniform illumination. (b) Deep learning methods typically incorporate priors such as Retinex into new architectures to improve performance, but they still struggle with complex lighting and may introduce artifacts or color inconsistencies. (c) In contrast, our method fully leverages intrinsic image priors to achieve more coherent and consistent enhancement.

As shown in Figure 1(a), early traditional LLIE methods primarily relied on handcrafted priors [Arici *et al.*, 2009; Lee *et al.*, 2013; Wang *et al.*, 2013] to enhance brightness by expanding dynamic range and contrast, and on Retinex-based models [Kimmel *et al.*, 2003; Guo *et al.*, 2016] that sought

*Corresponding author.
†Corresponding author.

to separate illumination from reflectance. Although effective in simple cases, these methods often struggled with complex degradations and exhibited limited generalization. With the advent of Deep Learning (DL), recent enhancement models [Zhang *et al.*, 2025a; Wu *et al.*, 2022; Yan *et al.*, 2025] have significantly improved restoration quality by learning mappings from low-light to normal-light images. For example, URetinexNet [Wu *et al.*, 2022] unfolds the Retinex decomposition process into a learnable deep network, enabling more stable estimation of illumination and reflectance. CIDNet [Yan *et al.*, 2025] introduces a new HVI color space specifically designed for low-light enhancement, allowing the network to more effectively disentangle brightness and structural information. Although these methods can enhance images effectively, many of them still overlook the intrinsic illumination priors inherent to low-light conditions, as illustrated in Figure 1(b).

To overcome these limitations, we introduce InterLight, a novel LLIE framework that systematically leverages intrinsic illumination priors. Unlike previous approaches that treat enhancement as generic image-to-image translation, InterLight builds an illumination-aware pipeline that injects sensor-level priors via physics-guided augmentation, represents degradations through adaptive prompts conditioned on latent illumination states, and employs a luminance-gated intrinsic memory to selectively compensate for information loss. The entire process is regularized by a self-supervised consistency objective, yielding natural colors, clearer textures, and visually coherent enhancements across diverse real-world scenarios.

In summary, our contributions are as follows:

- We introduce InterLight, a new LLIE framework that systematically exploits intrinsic image priors. It constructs an illumination-aware enhancement pipeline grounded in the physical and statistical properties of low-light images.
- We propose ICDE, a physics-guided augmentation strategy that simulates sensor-level illumination responses while preserving structural fidelity.
- We develop ADPG, which extracts sample-specific degradation prior through a learnable degradation dictionary. The resulting prompt provides both global and spatially adaptive guidance for the chrominance branch via the PRFB.
- We design a LGIM that retrieves learned structural and textural patterns to compensate for information loss. It adaptively strengthens enhancement in dark regions while preserving fidelity in well-lit areas.

2 Related Work

2.1 Low-Light Image Enhancement

Low-Light Image Enhancement (LLIE) [Wei *et al.*, 2018; Jiang *et al.*, 2024] aims to improve the quality and exposure of photographs captured in dark environments, producing dark-free images with natural color and proper illumination. Compared with traditional methods [Lee *et al.*, 2013; Wang *et al.*, 2013], current deep learning-based approaches primarily focus on minimizing the reconstruction error between the enhanced output and the ground truth. Their main progress lies in refining network architectures or learning strategies tailored for LLIE. For example, Zero-DCE [Guo *et al.*, 2020] estimates a reference curve via a deep network and performs zero-shot manner; PairLIE [Fu *et al.*, 2023] adopts an unsupervised framework that learns adaptive priors from paired low-light images; Retinexformer [Cai *et al.*, 2023] predicts illumination to brighten the image and then restores degradations. Other architectures, including normalizing flow models [Wang *et al.*, 2022], diffusion models [Jiang *et al.*, 2023], and Mamba-based models [Xu *et al.*, 2025b], have also been explored. In contrast to these approaches, we advocate fully leveraging intrinsic illumination priors for LLIE to achieve more coherent and self-consistent enhancement results.

2.2 Low-Light Image Enhancement with Intrinsic Priors

Early studies sought to improve LLIE performance by incorporating handcrafted priors [Arici *et al.*, 2009], such as the Dark Channel Prior or histogram-based constraints. These optimization-based methods explicitly embraced the idea of decomposition and laid the foundation for leveraging illumination priors. Following this trend, recent deep learning-based Retinex [Wei *et al.*, 2018; Yi *et al.*, 2025; Bai *et al.*, 2025] methods further extend this decomposition paradigm, yet their exploitation of illumination priors remains relatively shallow. Another line of work incorporates high-level visual priors [Zhang *et al.*, 2026b; Zhang *et al.*, 2026a; Xu *et al.*, 2023; Zhang *et al.*, 2025b], where auxiliary cues such as semantic segmentation or depth are used to guide the enhancement process. The intuition is that semantic structure may help regularize low-level restoration. However, these methods inevitably encounter a chicken-and-egg problem: reliable semantic or depth estimation itself requires sufficient visibility. Under complex real-world low-light conditions, such high-level predictions often become unstable or noisy, which can in turn introduce incorrect guidance and lead to artifacts or structural inconsistencies in the enhanced results. A third category of methods operates at the feature level [Xiao *et al.*, 2025; Xu *et al.*, 2025a; Li *et al.*, 2024; Zhao *et al.*, 2026; Zhou *et al.*, 2024; Chang *et al.*, 2026; Zhu *et al.*, 2026; Di *et al.*, 2025; He *et al.*, 2023; Yang *et al.*, 2022]. Some approaches employ codebooks or lookup tables that store representative features, enabling each input to query and retrieve suitable references for the decoder. In this work, we propose a hierarchical data–feature enhancement strategy fully exploiting intrinsic priors at both data-statistics and feature-representation levels. Besides, we explicitly incorporate degradation priors to guide adaptive restoration under diverse low-light conditions.

3 Method

In this paper, we propose InterLight, a novel LLIE framework that mines the intrinsic priors of low-light data without relying on external data or pre-trained priors. As illustrated in Figure 2, our framework integrates three synergistic mechanisms to address data scarcity and feature degradation: (1)

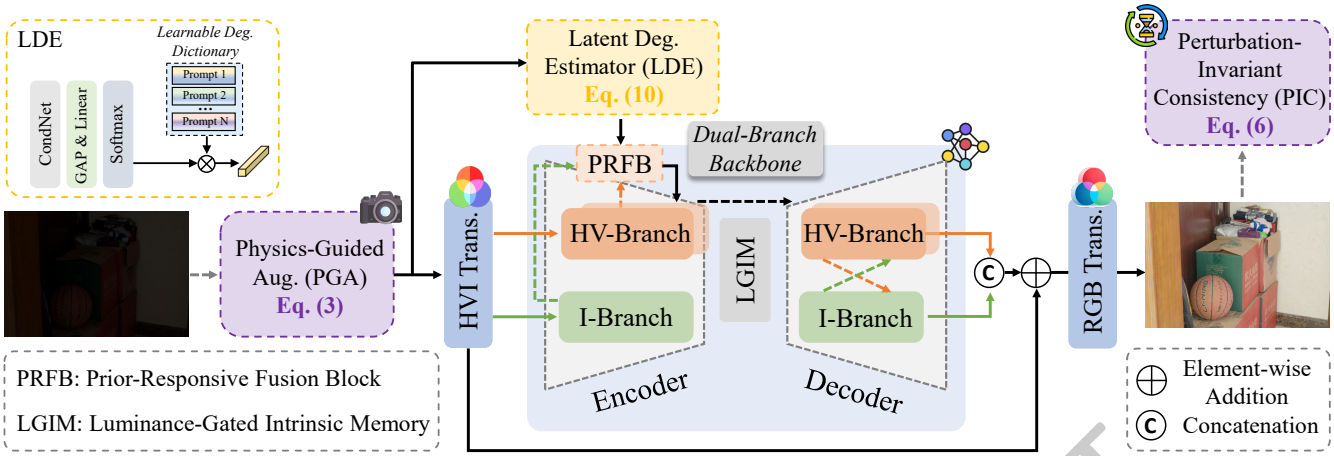

Figure 2: The overview of the proposed InterLight. The input is first augmented via PGA and then encoded by the LDE to produce an degradation prompt. After transforming the image into the HVI space, a dual-branch network restores illumination and chrominance with prompt-aware fusion. LGIM further compensates for information loss, and the final output is obtained through inverse HVI transformation with a residual connection.

Intrinsic-Consistent Data Expansion to simulate physically plausible domain shifts and enforce structural consistency; (2) Adaptive Degradation Prior Generation to inject image-specific degradation contexts; and (3) Luminance-Gated Intrinsic Memory to retrieve restoration knowledge based on local feature.

3.1 Overview

Figure 2 illustrates the complete pipeline of InterLight. During training, the input $\mathbf{I}_{in}$ first undergoes Physics-Guided Augmentation (PGA) to simulate sensor variations. The augmented image is then processed by the Latent Degradation Estimator (LDE) to extract a global context prompt $\mathbf{p}$ encoding the degradation condition. Next, the input is decomposed into the HVI color space, yielding chrominance channels $(\mathcal{H}, \mathcal{V})$ and intensity channel $\mathcal{I}$.

The disentangled representation is processed through a dual-branch U-Net architecture. One branch restores illumination from $\mathcal{I}$, focusing on brightness recovery, while the other refines chrominance with prompt guidance via PRFB, preserving color fidelity. Both branches employ four-level encoder-decoders with Lightweight Cross-Attention (LCA) for mutual information exchange. At the bottleneck, the Luminance-Gated Intrinsic Memory (LGIM) retrieves learned patterns to compensate information loss, with stronger enhancement applied to darker regions.

Finally, the decoded features are fused through inverse HVI transformation with a global residual connection to produce the output $\mathbf{I}_{out}$. During training, Perturbation-Invariant Consistency (PIC) further enforces output stability via self-supervised augmentation. PGA and PIC are bypassed during inference.

3.2 HVI Color Space

To effectively decouple illumination and reflectance information, we adopt the Horizontal/Vertical-Intensity (HVI) color space introduced in [Yan *et al.*, 2025]. Unlike standard color spaces, HVI incorporates a density-adaptive mechanism to handle the instability of chromaticity in low-light conditions.

Given an input image $\mathbf{I}_{in} \in \mathbb{R}^{H\times W\times 3}$, the intensity component is defined as the maximum value across channels, denoted as $\mathcal{P} = \max_c(\mathbf{I}^c_{in})$. To mitigate noise in dark regions, a density function $\mathbf{C}_k$ is computed based on the intensity:

$$\mathbf{C}_k = \left(\sin\left(\frac{\pi\mathcal{P}}{2}\right) + \epsilon\right)^k, \tag{1}$$

where k is a learnable parameter initialized to 0.2, and $\epsilon = 10^{-8}$ ensures numerical stability. As $\mathcal{P} \to 0$, $\mathbf{C}_k$ approaches zero, naturally suppressing unreliable color information.

Let S and H denote the standard saturation and normalized hue derived from the RGB-to-HSV conversion [Yan *et al.*, 2025]. The final HVI representation is formulated as:

$$\mathcal{H} = \mathbf{C}_k \cdot S \cdot \cos(2\pi H), \quad \mathcal{V} = \mathbf{C}_k \cdot S \cdot \sin(2\pi H), \quad \mathcal{I} = \mathcal{P}. \tag{2}$$

Here, $\mathcal{H}$ and $\mathcal{V}$ encode the chrominance information in a polar coordinate system modulated by $\mathbf{C}_k$, while $\mathcal{I}$ represents the illumination. The inverse transformation recovers the RGB image by normalizing $(\hat{\mathcal{H}}, \hat{\mathcal{V}})$ with $\mathbf{C}_k(\hat{\mathcal{I}})$ and applying the standard HSV-to-RGB conversion.

3.3 Intrinsic-Consistent Data Expansion

To overcome data scarcity without external datasets, we introduce ICDE with two physics-aware strategies.

Physics-Guided Augmentation (PGA). Standard augmentation often violates low-light physics by amplifying noise in dark regions. PGA simulates sensor response variations through mild channel-wise Gamma correction $\gamma_c \sim \mathcal{U}(0.95, 1.05)$ for each channel $c \in \{R, G, B\}$, where $\mathcal{U}(a, b)$ denotes the uniform distribution over $[a, b]$. An information protection mechanism preserves dark regions:

$$\mathbf{I}_{pga} = \alpha \cdot \mathbf{I}^{\gamma} + (1 - \alpha) \cdot \mathbf{I}, \tag{3}$$

$$\alpha = 3t^2 - 2t^3, \quad t = \min\left(1, \frac{\mathcal{P}}{\tau_d}\right), \tag{4}$$

where $\mathcal{P} = \max_c(\mathbf{I}_c)$ denotes the per-pixel maximum intensity across RGB channels, and $\tau_d = 0.05$ is the dark region threshold. The smoothstep function $\alpha \in [0, 1]$ ensures augmentation only affects regions with valid photon information ($\alpha \to 1$ when bright, $\alpha \to 0$ when dark), thereby avoiding non-physical artifacts in noise-dominated areas.

Perturbation-Invariant Consistency (PIC). Given an enhanced output $\mathbf{I}_e$, we generate weak and strong augmented views:

$$\mathbf{I}_w = \text{CenterCrop}_s(\mathbf{I}_e), \quad \mathbf{I}_s = \mathcal{G}_\sigma(\mathbf{I}_w), \tag{5}$$

where $s = 16$ is the crop size and $\mathcal{G}_\sigma$ denotes Gaussian blur with kernel size ranging from 9 to 21 and standard deviation $\sigma \sim \mathcal{U}(0.1, 5)$. The consistency loss minimizes the MSE distance between views:

$$\mathcal{L}_{consistency} = \beta(t) \cdot \|\mathbf{I}_w - \mathbf{I}_s\|_2^2, \tag{6}$$

where the weight $\beta(t)$ follows a cosine decay schedule to prevent over-smoothing as training converges:

$$\beta(t) = \frac{\beta_0}{2}\left(\cos\left(\frac{\pi t}{T}\right) + 1\right), \tag{7}$$

with $\beta_0 = 0.1$ being the initial weight and T the total training steps. This encourages learning of scale- and blur-invariant features intrinsic to the data.

3.4 Adaptive Degradation Prior Generation

While ICDE addresses data diversity, global degradation levels vary significantly across samples. The ADPG module extracts and injects degradation priors through two components.

Latent Degradation Estimator (LDE). Given the PGA output $\mathbf{I}_{pga} \in \mathbb{R}^{3\times H\times W}$, LDE first extracts a compact global feature vector using a lightweight condition network $\mathcal{C}$ (a stack of strided convolutions followed by 1×1 convolutions):

$$\mathbf{z} = \text{GAP}\left(\mathcal{C}(\mathbf{I}_{pga})\right) \in \mathbb{R}^{d_z}, \tag{8}$$

where GAP denotes global average pooling and $d_z = 32$ is the latent feature dimension. We maintain a learnable degradation dictionary $\mathbf{D} = [\mathbf{d}_1, \ldots, \mathbf{d}_K]^\top \in \mathbb{R}^{K\times d_p}$, where each vector $\mathbf{d}_k \in \mathbb{R}^{d_p}$ represents a prototypical degradation pattern. The coefficients over this dictionary are computed via softmax normalization:

$$\boldsymbol{\alpha} = \text{Softmax}\left(\mathbf{W}_\alpha \mathbf{z}\right) \in \mathbb{R}^K, \tag{9}$$

where $\mathbf{W}_\alpha \in \mathbb{R}^{K\times d_z}$ is a learnable projection matrix and $K = 32$ is the number of dictionary atoms. The context prompt $\mathbf{p}$ is obtained as a weighted combination of dictionary vectors:

$$\mathbf{p} = \phi\left(\boldsymbol{\alpha}^\top \mathbf{D}\right) = \phi\left(\sum_{k=1}^{K} \alpha_k \mathbf{d}_k\right) \in \mathbb{R}^{d_p}, \tag{10}$$

where $\phi(\cdot)$ denotes the GELU activation function and $d_p = 512$ is the prompt dimension. The dictionary vectors $\{\mathbf{d}_k\}$ are jointly learned during training, discovering representative degradation patterns from data without manual categorization.

Prior-Responsive Fusion Block (PRFB)

PRFB replaces standard cross-attention in the HV-branch encoder, modulating features using prompt $\mathbf{p}$. Specifically, $\mathbf{p}$ is first projected to channel-wise scale and shift parameters that adjust global feature statistics:

$$\boldsymbol{\gamma} = \sigma\left(\mathbf{W}_\gamma \mathbf{p}\right) \in \mathbb{R}^C, \quad \boldsymbol{\beta} = \mathbf{W}_\beta \mathbf{p} \in \mathbb{R}^C, \tag{11}$$

where $\sigma(\cdot)$ is the sigmoid function ensuring $\boldsymbol{\gamma} \in (0, 1)$, C denotes the feature channel dimension, and $\mathbf{W}_\gamma, \mathbf{W}_\beta \in \mathbb{R}^{C\times d_p}$ are learnable linear projections. The input feature $\mathbf{X} \in \mathbb{R}^{C\times H\times W}$ is then modulated as:

$$\mathbf{X}' = \mathbf{X} \odot (1 + \boldsymbol{\gamma}) + \boldsymbol{\beta}, \tag{12}$$

where $\odot$ denotes channel-wise multiplication with spatial broadcasting.

To enable spatially-adaptive guidance, the prompt is projected to a spatial feature map and upsampled:

$$\begin{aligned} \mathbf{P}_s &= \text{Upsample}_{H\times W}(\mathbf{U}) \in \mathbb{R}^{C\times H\times W}, \\ \mathbf{U} &= \text{Conv}\left(\text{Reshape}_{h\times w}\left(\mathbf{W}_s \mathbf{p}\right)\right), \end{aligned} \tag{13}$$

where $\mathbf{W}_s \in \mathbb{R}^{(C\cdot h\cdot w)\times d_p}$ projects the prompt to spatial dimensions, and h, w are level-dependent spatial sizes (16, 8, 4 for shallow, middle, deep layers respectively). A learned gate $\mathbf{G}$ balances degradation guidance versus local content:

$$\mathbf{G} = \sigma\left(\text{Conv}_{1\times1}\left(\text{DWConv}\left([\mathbf{q}_1; \mathbf{q}_2]\right)\right)\right) \in \mathbb{R}^{C\times H\times W}, \tag{14}$$

where $\mathbf{q}_1 = \text{Conv}_{1\times1}(\mathbf{P}_s)$, $\mathbf{q}_2 = \text{Conv}_{1\times1}(\mathbf{X}')$ are query projections, $[\cdot;\cdot]$ denotes channel-wise concatenation, and DWConv denotes depth-wise convolution. The query for cross-attention is formulated as:

$$\mathbf{Q} = \text{DWConv}\left(\mathbf{G} \odot \mathbf{q}_1 + (1 - \mathbf{G}) \odot \mathbf{q}_2\right). \tag{15}$$

The cross-attention with the I-branch feature $\mathbf{Y} \in \mathbb{R}^{C\times H\times W}$ follows normalized attention:

$$\mathbf{K}, \mathbf{V} = \text{Split}\left(\text{DWConv}\left(\text{Conv}_{1\times1}(\mathbf{Y})\right)\right), \tag{16}$$

$$\text{Attn}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax}\left(\frac{\bar{\mathbf{Q}}\bar{\mathbf{K}}^\top}{\tau_a}\right)\mathbf{V}. \tag{17}$$

The final output combines the attention result with a feed-forward network (FFN):

$$\mathbf{X}_{out} = \mathbf{X} + \text{Attn}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) + \text{FFN}(\mathbf{X}). \tag{18}$$

This design enables the network to adapt its restoration behavior based on the estimated global illumination condition, with the gate $\mathbf{G}$ learning to emphasize prompt guidance in severely degraded regions while preserving local content in well-lit areas.

3.5 Luminance-Gated Intrinsic Memory

Low-light images often suffer from information loss that local convolutions cannot recover. LGIM builds an internal knowledge base to compensate.

Specifically, LGIM maintains L learnable global vectors $\{\mathbf{m}_l^v\}_{l=1}^L \in \mathbb{R}^C$ designed to capture channel statistics alongside local patches $\{\mathbf{m}_l^p\}_{l=1}^L \in \mathbb{R}^{C\times r\times r}$ with patch size r to capture textures. These entries are dynamically updated

during training through sample-wise incorporation and mutual propagation. Given an input feature $\mathbf{F}_{in} \in \mathbb{R}^{C\times H\times W}$, the system queries this memory bank via attention mechanisms to retrieve relevant patterns at both vector and patch granularities, producing the retrieved memory feature denoted as $\mathbf{F}_{mem} \in \mathbb{R}^{C\times H\times W}$. Based on the critical observation that bright regions usually contain sufficient information whereas dark regions require aggressive hallucination, we implement an intensity-inverse modulation strategy. Unlike standard memory banks that execute spatially uniform feature retrieval, our LGIM adaptively bypasses memory intervention in high-fidelity regions while aggressively compensating for structural loss in dark areas. The final feature fusion is thus formulated as:

$$\mathbf{F}_{out} = \mathbf{F}_{in} + \lambda \cdot (1 + \sigma(\eta)(1 - g)) \cdot \mathbf{F}_{mem}, \tag{19}$$

where $g = \sigma(\text{MLP}(\text{GAP}(\mathbf{F}_{in}))) \in [0, 1]$ is a learned gate indicating local brightness estimated from input feature statistics, λ is a learnable fusion scale, η is a learnable adaptive strength parameter, and $\sigma(\cdot)$ denotes the sigmoid function. When $g \to 0$ (dark), the adaptive gain $(1 + \sigma(\eta)(1 - g))$ increases; when $g \to 1$ (bright), it decreases toward 1. This ensures aggressive memory retrieval for degraded dark regions while preserving fidelity in well-lit areas.

The I-branch uses stronger fusion ($\lambda_{init} = 1.2$) with an additional learnable brightness bias $\mathbf{b} \in \mathbb{R}^C$, while the HV-branch uses conservative fusion ($\lambda_{init} = 0.8$) with I-branch features as brightness reference for cross-branch guidance.

3.6 Training Objectives

The network is supervised by a compound loss in both RGB and HVI domains:

$$\mathcal{L}_{rec} = \mathcal{L}_{L1} + \mathcal{L}_{ssim} + \mathcal{L}_{edge} + \mu_p\mathcal{L}_{perc}, \tag{20}$$

combining pixel-wise L1 loss, structural SSIM loss, Laplacian edge loss, and VGG-based perceptual loss. The HVI-domain loss (weighted by μ_{hvi}) ensures accurate reconstruction of both illumination and reflectance components:

$$\mathcal{L}_{total} = \mathcal{L}^{RGB}_{rec} + \mu_{hvi}\mathcal{L}^{HVI}_{rec} + \mathcal{L}_{consistency}. \tag{21}$$

A dual-path training strategy supervises both baseline output $\hat{\mathbf{I}}_{base}$ (without LGIM) and memory-enhanced output $\hat{\mathbf{I}}_{mem}$:

$$\mathcal{L}_{dual} = \mathcal{L}_{total}(\hat{\mathbf{I}}_{base}) + \lambda_{lgim}\mathcal{L}_{total}(\hat{\mathbf{I}}_{mem}), \tag{22}$$

ensuring robust baseline performance while enabling complementary memory-based enhancement.

4 Experiments

4.1 Experimental Setup

Datasets and Metrics. Following established protocols in low-light image enhancement, we conduct comprehensive evaluations across multiple benchmarks. We adopt the LOL-v1 [Wei *et al.*, 2018], LOL-v2 [Yang *et al.*, 2021] , SICE [Cai *et al.*, 2018] (including the Mix and Grad test sets, SID (Sony-Total-Dark) [Chen *et al.*, 2018], and LSRW-Huawei [Hai *et al.*, 2023] datasets. We employ widely adopted metrics, including Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) [Wang *et al.*, 2004], to quantify restoration fidelity. Higher scores indicate better performance.

Implementation Details. The training process is optimized using the Adam [Kinga *et al.*, 2015] optimizer with $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The initial learning rate is set to 2×10^{-4} and decays following a cosine annealing schedule. We train the model for 1500 epochs with a batch size of 8. Notably, InterLight consistently converges across four fundamentally distinct datasets (LOL, SID, SICE, LSRW) using this unified optimization setting, without requiring any per-dataset initialization tuning, thereby verifying its excellent optimization stability. The input images are randomly cropped to 256×256 patches and augmented with random horizontal and vertical flips. The dual-branch encoder-decoder follows a four-level U-Net architecture with channel dimensions $[36, 36, 72, 144]$. All experiments are conducted on NVIDIA RTX 4090 GPUs using PyTorch.

For the ADPG module, we employ a degradation dictionary with $N_a = 32$ and prompt dimension $d_p = 512$. The spatial prior resolution in PRFB is set to $\{16, 8, 4\}$ for the three encoder levels. For the LGIM module, we maintain $L = 16$ memory entries with patch size $k = 4$, and initialize the fusion scales to $\lambda_I = 1.2$ for the I-branch and $\lambda_{HV} = 0.8$ for the HV-branch. The loss coefficients are set as follows: HVI-domain weight $\mu_{hvi} = 0.5$, perceptual loss weight $\mu_p = 0.1$, and LGIM dual-path weight $\lambda_{lgim} = 1.0$. The PIC consistency loss weight β_0 is initialized to 0.1 and decays via cosine annealing. For PGA, we apply channel-wise Gamma correction with $\gamma \sim \mathcal{U}(0.95, 1.05)$ at probability $p = 0.3$. The dark protection threshold is set to $\tau = 0.05$.

4.2 Comparison with State-of-the-Art Methods

We compare InterLight with a wide range of state-of-the-art methods, including Retinex-based methods (RetinexNet [Wei *et al.*, 2018], KinD [Zhang *et al.*, 2019], RUAS [Liu *et al.*, 2021]), RGB-based methods (Zero-DCE [Guo *et al.*, 2020], EnlightenGAN [Jiang *et al.*, 2021], and LLFormer [Wang *et al.*, 2023b]), and recent advanced models (Retinexformer [Cai *et al.*, 2023], DarkIR [Feijoo *et al.*, 2025], CIDNet [Yan *et al.*, 2025], and CWNet [Zhang *et al.*, 2025a]).

Results on LOL Datasets

Table 1 reports the quantitative results on LOL-v1 and LOL-v2 datasets. Our InterLight achieves consistent superiority across all subsets. On the LOL-v1 dataset, InterLight achieves the second-best PSNR of 24.78 dB and the best SSIM of 0.862. Notably, compared to the recent HVI-based method CIDNet, our method yields a significant improvement of 0.97 dB in PSNR, validating the effectiveness of our intrinsic learning paradigm over standard HVI transformations. On the LOL-v2-Real subset, InterLight ranks first with 24.06 dB PSNR and 0.866 SSIM, demonstrating robustness to complex real-world noise. On the LOL-v2-Synthetic, it also secures the top rank with 25.73 dB PSNR, significantly outperforming LightenDiff at much lower computational cost. The visual results in Figure 3 show that our method not only provides more accurate recovery of multi-color regions than the CIDNet, but also delivers more stable brightness enhancement.

Methods	Venue	Color Space	Complexity		LOL-v1		LOL-v2-Real		LOL-v2-Syn	
			Params/M	FLOPs/G	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑
RetinexNet [Wei *et al.*, 2018]	BMVC'18	Retinex	0.84	584.47	18.92	0.427	16.10	0.401	17.14	0.762
KinD [Zhang *et al.*, 2019]	MM'19	Retinex	8.02	34.99	23.02	0.843	17.54	0.669	18.32	0.796
ZeroDCE [Guo *et al.*, 2020]	CVPR'20	RGB	0.075	4.83	21.88	0.640	16.06	0.580	17.71	0.815
RUAS [Liu *et al.*, 2021]	CVPR'21	Retinex	0.003	0.83	18.65	0.518	15.33	0.488	13.77	0.638
EnlightenGAN [Jiang *et al.*, 2021]	TIP'21	RGB	114.35	61.01	20.00	0.691	18.23	0.617	16.57	0.734
SNR-Aware [Xu *et al.*, 2022]	CVPR'22	SNR+RGB	4.01	26.35	24.61	0.842	21.48	0.849	24.14	0.928
Bread [Guo and Hu, 2023]	IJCV'23	YCbCr	2.02	19.85	22.92	0.836	20.83	0.847	17.63	0.919
PairLIE [Fu *et al.*, 2023]	CVPR'23	Retinex	0.33	20.81	23.53	0.755	19.89	0.778	19.07	0.794
LLFormer [Wang *et al.*, 2023b]	AAAI'23	RGB	24.55	22.52	23.65	0.816	20.06	0.792	24.04	0.909
Retinexformer [Cai *et al.*, 2023]	ICCV'23	Retinex	1.53	15.85	25.16	0.845	22.79	0.840	25.67	0.930
LightenDiff [Jiang *et al.*, 2024]	ECCV'24	RGB	26.54	2257.42	23.62	0.829	22.88	0.855	21.58	0.869
CWNet [Zhang *et al.*, 2025a]	ICCV'25	RGB	1.23	11.3	23.60	0.850	21.65	0.860	25.50	0.936
CIDNet [Yan *et al.*, 2025]	CVPR'25	HVI	1.88	7.57	23.81	0.857	23.90	0.865	25.71	0.942
Ours	-	HVI	10.91	8.41	24.78	0.862	24.06	0.866	25.73	0.935

Table 1: Quantitative results of PSNR↑ and SSIM↑ on the LOL (v1 and v2) datasets. The best performance is in red color and the second best is in blue color.

Figure 3: Visual comparison of enhanced results from SOTA methods on LOL-v1 (top) and LOL-v2-Real (bottom).

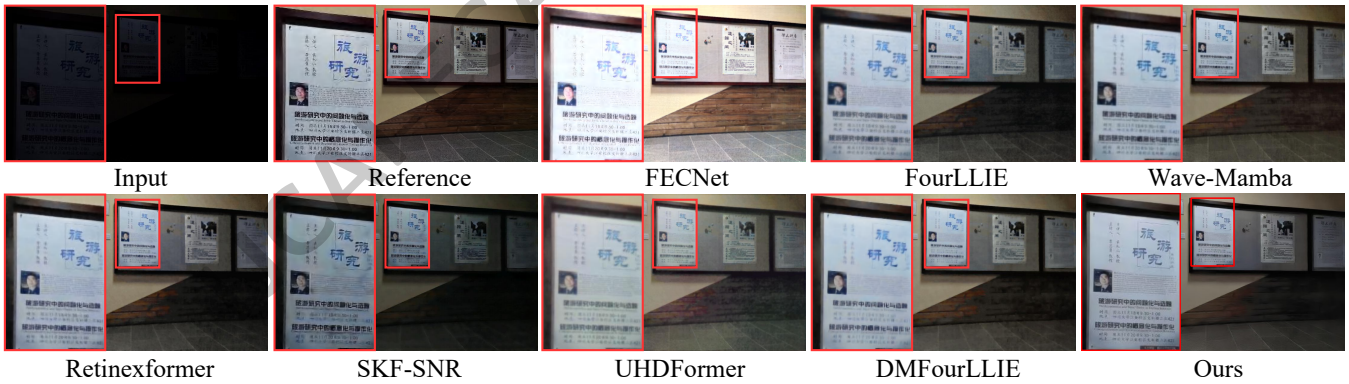

Figure 4: Visual comparison on LSRW-Huawei dataset.

Results on Results on SICE and SID Datasets

To evaluate the capability of handling extreme degradations, we test on the SICE (multi-exposure) and SID (extreme low-light) datasets. As shown in Table 2, InterLight consistently outperforms existing methods. Specifically, on the SID dataset, which presents extreme non-linear sensor noise and severe raw-to-RGB color shifts, our method achieves 22.98 dB PSNR, surpassing the strong competitor CIDNet by 0.08 dB and the flow-based LLFlow by over 6.7 dB. This proves our resilience against severe non-physical degradations, as the dynamic projection in the ADPG module effectively captures sample-specific non-linear prior contexts. Additionally,

Methods	SICE		SID	
	PSNR↑	SSIM↑	PSNR↑	SSIM↑
RetinexNet [Wei *et al.*, 2018]	12.42	0.613	15.70	0.395
ZeroDCE [Guo *et al.*, 2020]	12.45	0.639	14.09	0.090
RUAS [Liu *et al.*, 2021]	8.66	0.494	12.62	0.081
URetinexNet [Wu *et al.*, 2022]	10.90	0.605	15.52	0.323
LLFlow [Wang *et al.*, 2022]	12.74	0.617	16.23	0.367
CIDNet [Yan *et al.*, 2025]	13.44	0.642	22.90	0.676
Ours	13.56	0.653	22.98	0.681

Table 2: Quantitative result on SICE and Sony-Total-Dark datasets. The top-ranking score is in red.

Methods	Venue	LSRW-Huawei	
		PSNR↑	SSIM↑
KinD [Zhang *et al.*, 2019]	MM'19	16.58	0.569
SNR-Aware [Xu *et al.*, 2022]	CVPR'22	20.67	0.591
Retinexformer [Cai *et al.*, 2023]	ICCV'23	21.23	0.631
FourLLIE [Wang *et al.*, 2023a]	MM'23	21.11	0.626
Wave-Mamba [Zou *et al.*, 2024]	MM'24	21.19	0.631
DMFourLLIE [Zhang *et al.*, 2024]	MM'24	21.09	0.633
DarkIR [Feijoo *et al.*, 2025]	CVPR'25	18.93	0.583
Ours	-	21.39	0.625

Table 3: Quantitative result on LSRW-Huawei datasets. The top-ranking score is in red and the second best is in blue color.

this highlights the advantage of our Luminance-Gated Intrinsic Memory (LGIM) in hallucinating valid details from extremely weak signals where local convolutions often fail. Similarly, on the SICE dataset, we achieve the best performance (13.56 dB), demonstrating robust dynamic range compression capabilities.

Results on LSRW-Huawei dataset

We further evaluate InterLight on the LSRW-Huawei dataset, which involves images captured with sensor characteristics distinct from those in the LOL dataset. As shown in Table 3, InterLight achieves the best PSNR of 21.39 dB, surpassing recent approaches such as DMFourLLIE and DarkIR. These results confirm that our Physics-Guided Augmentation (PGA) strategy effectively adapts internal feature distributions to unseen sensor domains. As shown in Figure 4, while InterLight does not surpass DMFourLLIE in the SSIM metric in Table 3, its enhancement of character regions appears more realistic. This may be attributed to our method's stronger ability to recover fine details in dark areas.

4.3 Ablation Study

To validate the contribution of each component in InterLight, we conduct a comprehensive ablation study on the LOL-v1 dataset. We evaluate the impact of the Adaptive Degradation Prior Generation (ADPG) framework, Luminance-Gated Intrinsic Memory (LGIM), and Intrinsic-Consistent Data Expansion (ICDE).

As shown in Table 4, the baseline model yields a PSNR of 23.46 dB. Incorporating ADPG (Index b) brings a substantial gain of 0.75 dB, demonstrating that extracting image-specific degradation priors is crucial for adaptive restoration.

Index	ADPG	LGIM	ICDE	PSNR↑	SSIM↑	ΔPSNR
a	✗	✗	✗	23.46	0.842	–
b	✓	✗	✗	24.21	0.859	+0.75
c	✗	✓	✗	24.27	0.859	+0.81
d	✗	✗	✓	23.87	0.848	+0.41
e	✓	✓	✗	23.94	0.855	+0.48
f	✓	✗	✓	24.32	0.856	+0.86
g	✗	✓	✓	24.16	0.855	+0.70
Ours	✓	✓	✓	**24.78**	**0.866**	**+1.32**

Table 4: Ablation study on the LOL-v1 dataset. We evaluate all combinations of three proposed components: ADPG, LGIM, and ICDE. Δ denotes the PSNR improvement over the vanilla baseline.

Furthermore, adding LGIM alone (Index c) improves performance by 0.81 dB, which validates our hypothesis that retrieving high-quality features from an internal memory bank effectively compensates for information loss in dark regions. Finally, employing ICDE (Index d) leads to a 0.41 dB improvement, confirming that our physics-aware augmentation facilitates the learning of more robust features.

Finally, the full model, synergizing all three components, achieves the peak performance of 24.78 dB, with a cumulative improvement of 1.32 dB over the baseline.Crucially, while the ICDE strategy provides an orthogonal improvement (+0.41 dB), our architectural modules deliver the vast majority of the gains (ADPG and LGIM independently contribute +0.75 dB and +0.81 dB, respectively). This confirms that our intrinsic structural design is the definitive core contributor, and all proposed modules are complementary and collectively contribute to the state-of-the-art performance of InterLight.

5 Conclusion and Future Work

In this paper, we propose InterLight, a principled LLIE framework that deeply mines intrinsic illumination priors at both the data and feature levels. Through Intrinsic-Consistent Data Expansion, the model learns illumination-invariant representations under physically plausible perturbations. The Adaptive Degradation Prior Generation module captures sample-specific degradation characteristics via a learnable dictionary, producing prompts that guide chrominance restoration in a spatially adaptive manner. Luminance-Gated Intrinsic Memory further retrieves structural and textural cues with luminance-aware modulation, enabling stronger compensation in dark areas while maintaining fidelity in bright regions. Extensive experiments across diverse benchmarks validate the effectiveness of InterLight.

Despite its strong performance, the dual-branch architecture with LGIM introduces additional computational overhead, and the physics-guided augmentation assumes relatively linear degradation models. In future work, we plan to explore model compression for lightweight deployment and extend this paradigm to video low-light enhancement by leveraging temporal coherence.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant 62302105, and in part by Funding by Science and Technology Projects in Guangzhou under Grant 2025A04J3851.

References

[Arici *et al.*, 2009] Tarik Arici, Salih Dikbas, and Yucel Altunbasak. A histogram modification framework and its application for image contrast enhancement. *IEEE TIP*, 18(9):1921–1935, 2009.

[Bai *et al.*, 2025] Haowen Bai, Jiangshe Zhang, Zixiang Zhao, Lilun Deng, Yukun Cui, and Shuang Xu. Retinexmef: Retinex-based glare effects aware unsupervised multi-exposure image fusion. In *ICCV*, 2025.

[Cai *et al.*, 2018] Jianrui Cai, Shuhang Gu, and Lei Zhang. Learning a deep single image contrast enhancer from multi-exposure images. *IEEE TIP*, 27(4):2049–2062, 2018.

[Cai *et al.*, 2023] Yuanhao Cai, Hao Bian, Jing Lin, Haoqian Wang, Radu Timofte, and Yulun Zhang. Retinexformer: One-stage retinex-based transformer for low-light image enhancement. In *ICCV*, pages 12504–12513, 2023.

[Chang *et al.*, 2026] Laibin Chang, Yunke Wang, Bo Du, and Chang Xu. Color correction meets cross-spectral refinement: a distribution-aware diffusion for underwater image restoration. *IEEE TMM*, pages 1–14, 2026.

[Chen *et al.*, 2018] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *CVPR*, 2018.

[Di *et al.*, 2025] Xin Di, Long Peng, Peizhe Xia, Wenbo Li, Renjing Pei, Yang Cao, Yang Wang, and Zheng-Jun Zha. Qmambabsr: Burst image super-resolution with query state space model. In *CVPR*, 2025.

[Feijoo *et al.*, 2025] Daniel Feijoo, Juan C. Benito, Alvaro Garcia, and Marcos V. Conde. Darkir: Robust low-light image restoration. In *CVPR*, pages 10879–10889, 2025.

[Fu *et al.*, 2023] Zhenqi Fu, Yan Yang, Xiaotong Tu, Yue Huang, Xinghao Ding, and Kai-Kuang Ma. Learning a simple low-light image enhancer from paired low-light instances. In *CVPR*, pages 22252–22261, 2023.

[Gu *et al.*, 2025] Yuxuan Gu, Haoxuan Wang, Pengyang Ling, Zhixiang Wei, Huaian Chen, Yi Jin, and Enhong Chen. Improving visual and downstream performance of low-light enhancer with vision foundation models collaboration. In *CVPR*, pages 16071–16080, 2025.

[Guo and Hu, 2023] Xiaojie Guo and Qiming Hu. Low-light image enhancement via breaking down the darkness. *IJCV*, 131(1):48–66, Jan 2023.

[Guo *et al.*, 2016] Xiaojie Guo, Yu Li, and Haibin Ling. Lime: Low-light image enhancement via illumination map estimation. *IEEE TIP*, 26(2):982–993, 2016.

[Guo *et al.*, 2020] Chunle Guo 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 *CVPR*, 2020.

[Hai *et al.*, 2023] Jiang Hai, Zhu Xuan, Ren Yang, Yutong Hao, Fengzhu Zou, Fang Lin, and Songchen Han. R2rnet: Low-light image enhancement via real-low to real-normal network. *JVCI*, 90:103712, 2023.

[He *et al.*, 2023] Xiao He, Chang Tang, Xin Zou, and Wei Zhang. Multispectral object detection via cross-modal conflict-aware learning. In *ACM MM*, 2023.

[Jiang *et al.*, 2021] Yifan Jiang, Xinyu Gong, Ding Liu, Yu Cheng, Chen Fang, Xiaohui Shen, Jianchao Yang, Pan Zhou, and Zhangyang Wang. Enlightengan: Deep light enhancement without paired supervision. *IEEE TIP*, 30:2340–2349, 2021.

[Jiang *et al.*, 2023] Hai Jiang, Ao Luo, Haoqiang Fan, Songchen Han, and Shuaicheng Liu. Low-light image enhancement with wavelet-based diffusion models. *ACM TOG*, 42(6):1–14, 2023.

[Jiang *et al.*, 2024] Hai Jiang, Ao Luo, Xiaohong Liu, Songchen Han, and Shuaicheng Liu. Lightendiffusion: Unsupervised low-light image enhancement with latent-retinex diffusion models. In *ECCV*, 2024.

[Kimmel *et al.*, 2003] Ron Kimmel, Michael Elad, Doron Shaked, Renato Keshet, and Irwin Sobel. A variational framework for retinex. *IJCV*, 52(1):7–23, 2003.

[Kinga *et al.*, 2015] Diederik Kinga, Jimmy Ba Adam, et al. A method for stochastic optimization. In *ICLR*, 2015.

[Lee *et al.*, 2013] Chulwoo Lee, Chul Lee, and Chang-Su Kim. Contrast enhancement based on layered difference representation of 2d histograms. *IEEE TIP*, 22(12):5372–5384, 2013.

[Li *et al.*, 2018] Mading Li, Jiaying Liu, Wenhan Yang, Xiaoyan Sun, and Zongming Guo. Structure-revealing low-light image enhancement via robust retinex model. *IEEE TIP*, 27(6):2828–2841, 2018.

[Li *et al.*, 2024] Zhuoyuan Li, Zikun Yuan, Li Li, Dong Liu, Xiaohu Tang, and Feng Wu. Object segmentation-assisted inter prediction for versatile video coding. *IEEE TBC*, 70(4):1236–1253, 2024.

[Liu *et al.*, 2021] Risheng Liu, Long Ma, Jiaao Zhang, Xin Fan, and Zhongxuan Luo. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. In *CVPR*, pages 10561–10570, 2021.

[Ma *et al.*, 2022] Long Ma, Tengyu Ma, Risheng Liu, Xin Fan, and Zhongxuan Luo. Toward fast, flexible, and robust low-light image enhancement. In *CVPR*, 2022.

[Ma *et al.*, 2025] Long Ma, Tengyu Ma, Chengpei Xu, Jinyuan Liu, Xin Fan, Zhongxuan Luo, and Risheng Liu. Learning with self-calibrator for fast and robust low-light image enhancement. *IEEE TPAMI*, 47(10):9095–9112, 2025.

[Wang *et al.*, 2004] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE TIP*, 13(4):600–612, 2004.

[Wang *et al.*, 2013] Shuhang Wang, Jin Zheng, Hai-Miao Hu, and Bo Li. Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE TIP*, 22(9):3538–3548, 2013.

[Wang *et al.*, 2022] Yufei Wang, Renjie Wan, Wenhan Yang, Haoliang Li, Lap-Pui Chau, and Alex Kot. Low-light image enhancement with normalizing flow. In *AAAI*, 2022.

[Wang *et al.*, 2023a] Chenxi Wang, Hongjun Wu, and Zhi Jin. Fourllie: Boosting low-light image enhancement by fourier frequency information. In *ACM MM*, 2023.

[Wang *et al.*, 2023b] Tao Wang, Kaihao Zhang, Tianrun Shen, Wenhan Luo, Bjorn Stenger, and Tong Lu. Ultra-high-definition low-light image enhancement: A benchmark and transformer-based method. In *AAAI*, 2023.

[Wang *et al.*, 2025] Sen Wang, Shao Zeng, Tianjun Gu, Zhizhong Zhang, Ruixin Zhang, Shouhong Ding, Jingyun Zhang, Jun Wang, Xin Tan, Yuan Xie, et al. From enhancement to understanding: Build a generalized bridge for low-light vision via semantically consistent unsupervised fine-tuning. In *ICCV*, pages 13804–13814, 2025.

[Wei *et al.*, 2018] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. In *BMVC*, 2018.

[Wu *et al.*, 2022] Wenhui Wu, Jian Weng, Pingping Zhang, Xu Wang, Wenhan Yang, and Jianmin Jiang. Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement. In *CVPR*, pages 5901–5910, 2022.

[Wu *et al.*, 2025] Yuan Wu, Zhiqiang Yan, Yigong Zhang, Xiang Li, and Jian Yang. See through the dark: Learning illumination-affined representations for nighttime occupancy prediction. In *NeurIPS*, 2025.

[Xiao *et al.*, 2025] Zeyu Xiao, Zhuoyuan Li, and Wei Jia. Occlusion-embedded hybrid transformer for light field super-resolution. In *AAAI*, pages 8700–8708, 2025.

[Xu *et al.*, 2022] Xiaogang Xu, Ruixing Wang, Chi-Wing Fu, and Jiaya Jia. Snr-aware low-light image enhancement. In *CVPR*, pages 17693–17703, 2022.

[Xu *et al.*, 2023] Xiaogang Xu, Ruixing Wang, and Jiangbo Lu. Low-light image enhancement via structure modeling and guidance. In *CVPR*, pages 9893–9903, 2023.

[Xu *et al.*, 2025a] Xiaogang Xu, Jiafei Wu, Qingsen Yan, Jiequan Cui, Richang Hong, and Bei Yu. Learnable feature patches and vectors for boosting low-light image enhancement without external knowledge. In *ICCV*, 2025.

[Xu *et al.*, 2025b] Yichu Xu, Di Wang, Lefei Zhang, and Liangpei Zhang. Dual selective fusion transformer network for hyperspectral image classification. *NN*, 187:107311, 2025.

[Yan *et al.*, 2025] Qingsen Yan, Yixu Feng, Cheng Zhang, Guansong Pang, Kangbiao Shi, Peng Wu, Wei Dong, Jinqiu Sun, and Yanning Zhang. Hvi: A new color space for low-light image enhancement. In *CVPR*, 2025.

[Yang *et al.*, 2021] Wenhan Yang, Wenjing Wang, Haofeng Huang, Shiqi Wang, and Jiaying Liu. Sparse gradient regularized deep retinex network for robust low-light image enhancement. *IEEE TIP*, 30:2072–2086, 2021.

[Yang *et al.*, 2022] Canqian Yang, Meiguang Jin, Xu Jia, Yi Xu, and Ying Chen. Adaint: Learning adaptive intervals for 3d lookup tables on real-time image enhancement. In *CVPR*, pages 17522–17531, 2022.

[Yi *et al.*, 2025] Xunpeng Yi, Han Xu, Hao Zhang, Linfeng Tang, and Jiayi Ma. Diff-retinex++: Retinex-driven reinforced diffusion model for low-light image enhancement. *IEEE TPAMI*, 2025.

[Zhang *et al.*, 2019] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. Kindling the darkness: A practical low-light image enhancer. In *ACM MM*, pages 1632–1640, 2019.

[Zhang *et al.*, 2024] Tongshun Zhang, Pingping Liu, Ming Zhao, and Haotian Lv. Dmfourllie: Dual-stage and multi-branch fourier network for low-light image enhancement. In *ACM MM*, pages 7434–7443, 2024.

[Zhang *et al.*, 2025a] Tongshun Zhang, Pingping Liu, Yubing Lu, Mengen Cai, Zijian Zhang, Zhe Zhang, and Qiuzhan Zhou. Cwnet: Causal wavelet network for low-light image enhancement. In *ICCV*, 2025.

[Zhang *et al.*, 2025b] Xu Zhang, Huan Zhang, Guoli Wang, Qian Zhang, Lefei Zhang, and Bo Du. Uniuir: Considering underwater image restoration as an all-in-one learner. *IEEE TIP*, 34:6963–6977, 2025.

[Zhang *et al.*, 2026a] Xu Zhang, Jiaqi Ma, Guoli Wang, Qian Zhang, Huan Zhang, and Lefei Zhang. Perceive-ir: Learning to perceive degradation better for all-in-one image restoration. *IEEE TIP*, 35:2018–2033, 2026.

[Zhang *et al.*, 2026b] Xu Zhang, Huan Zhang, Guoli Wang, Qian Zhang, and Lefei Zhang. Clearair: A human-visual-perception-inspired all-in-one image restoration. In *AAAI*, pages 12861–12869, 2026.

[Zhao *et al.*, 2026] Rui Zhao, Ruiqin Xiong, Dongkai Wang, Shiyu Xuan, Jian Zhang, Xiaopeng Fan, and Tiejun Huang. Spike camera optical flow estimation based on continuous spike streams. *IEEE TPAMI*, 2026.

[Zhou *et al.*, 2024] Han Zhou, Wei Dong, Xiaohong Liu, Shuaicheng Liu, Xiongkuo Min, Guangtao Zhai, and Jun Chen. Glare: Low light image enhancement via generative latent feature based codebook retrieval. In *ECCV*, 2024.

[Zhu *et al.*, 2026] Linwei Zhu, Junhao Zhu, Xu Zhang, Huan Zhang, Ye Li, Runmin Cong, and Sam Kwong. Enhanced quality aware scalable underwater image compression. *ACM TOMM*, 2026.

[Zou *et al.*, 2024] Wenbin Zou, Hongxia Gao, Weipeng Yang, and Tongtong Liu. Wave-mamba: Wavelet state space model for ultra-high-definition low-light image enhancement. In *ACM MM*, pages 1534–1543, 2024.