






Region-Adaptive Low-Light Image Enhancement with Light Effect Suppression and Detail Preservation

Liheng Luo^{†,1,3} , Wantong Xie^{†,1} , Xiushan Xia¹ , Zerui Li^{3,4} , Yunbo Zhao^{‡,2,3} 

¹Institute of Advanced Technology, University of Science and Technology of China, Hefei, China

²Department of Automation, University of Science and Technology of China, Hefei, China

³Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, China

⁴Hefei Zhongke Liheng Intelligent Technology Co., Ltd., Hefei, China

Abstract

Low-light image enhancement seeks to improve the visual quality of images captured under poor illumination, yet existing methods often struggle with unnatural artifacts, overexposure, or detail loss, particularly in challenging real-world scenarios like underground coal mines. We propose a novel unsupervised region-adaptive framework that integrates light effect suppression and detail preservation to address these issues. Leveraging Retinex theory, our approach decomposes images into illumination and reflectance components, employing a region segmentation module to distinguish dark and bright areas for targeted enhancement. A lightweight denoising network mitigates noise, while an adaptive illumination enhancer and light effect suppressor collaboratively optimize illumination to ensure natural appearance and correct visual imbalances. A composite loss function balances brightness enhancement, structural integrity, and artifact suppression across regions. Extensive experiments on the LOL-v2, LSRW and our private datasets demonstrate superior performance. For instance, on our dataset, improvements of 3.26% in BRISQUE, 0.24% in NIQE, and 11.22% in PIQE were achieved compared to state-of-the-art methods, providing visually pleasing results with enhanced brightness, reduced artifacts, and preserved textures, making it well-suited for real-world applications.

CCS Concepts

• **Computing methodologies** → **Image processing; Artificial intelligence;**

1. Introduction

Low-light image enhancement is a fundamental problem in computer vision, aimed at improving the visibility, contrast, and overall quality of images captured in poor lighting conditions [JB25]. Such images are prevalent in real-world scenarios, including night photography, surveillance, and autonomous driving, where insufficient illumination leads to low contrast, high noise, and loss of critical details [SY24]. These degradations not only impair human perception but also hinder the performance of downstream tasks such as object detection, recognition, and scene understanding [XHZ*24].

Early traditional methods, such as histogram equalization [CS04] and gamma correction [HCC12], directly enhance visibility through simple transformations but largely ignore illumination factors, leading to visually unrealistic results. To address this, Retinex-based methods decompose images into reflectance and illumination components for adaptive enhancement [LSLH13]. Recent advances include RetinexDIP [ZXW*21], which leverages deep image priors,

and Diff-Retinex [YXZ*23], which integrates diffusion models, yet these remain computationally demanding and sensitive to initial estimates. In parallel, learning-based models have achieved significant breakthroughs through end-to-end deep neural networks. ZeroDCE [GLG*20] pioneered zero-reference enhancement via adaptive curve adjustments, while SCI [MML*22] employs self-calibrated illumination for exposure stability. EnlightenGAN [JGL*21] overcomes paired data constraints through global-local discriminators and attention mechanisms. Despite these innovations, existing methods frequently produce unnatural artifacts—such as halos around light sources or over-enhanced bright regions—and struggle to retain fine textures in complex scenes. While some existing methods incorporate region-aware processing strategies, they exhibit significant limitations when addressing the joint challenges of light effect suppression and detail preservation. LIME [Guo16] lacks region-specific processing, leading to local over-exposure in high dynamic range scenes. SICE [CGZ18] applies adaptive enhancement but does not address light effect suppression, resulting in halo artifacts. RUAS [LMZ*21] focus on computational efficiency but inadequately handle the complex interplay between dark region enhancement and bright region artifact suppression.

To address these limitations, we propose a novel region-adaptive

[†] These authors contributed equally to this work.

[‡] Corresponding author.

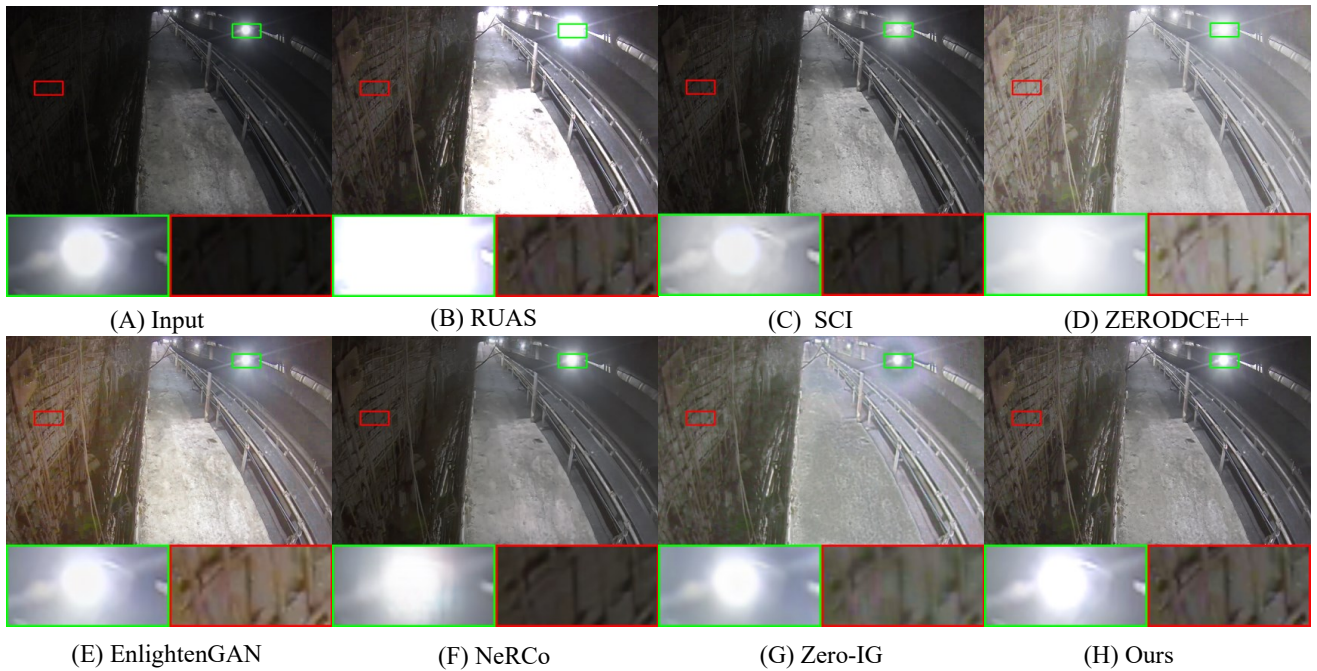


Figure 1: Visual comparison on the underground coal mine dataset. From left to right: input image, results from RUAS, SCI, ZeroDCE++, EnlightenGAN, NeRCO, Zero-IG, and our method, showcasing brightness enhancement and artifact suppression.

low-light image enhancement framework that integrates Retinex-based decomposition with sophisticated region-specific processing and light effect suppression strategies. Our approach employs a region segmentation module to distinguish dark and bright areas, enabling targeted illumination adjustments. A lightweight denoising network mitigates noise interference, while an adaptive illumination enhancer and light effect suppressor collaboratively optimize the illumination layer to achieve a natural appearance. A carefully designed composite loss function ensures end-to-end optimization, balancing brightness enhancement, structural integrity, and artifact suppression. Our main contributions are:

- A novel region-adaptive enhancement framework that leverages pixel-level segmentation to apply tailored processing strategies for dark and bright regions, uniquely integrating light effect suppression with detail preservation in a unified end-to-end architecture.
- A sophisticated light effect suppressor with attention-guided mechanism that specifically targets halo artifacts and over-enhancement in bright regions, addressing limitations not handled by existing adaptive methods.
- An innovative composite loss function incorporating region-specific reconstruction, structural consistency, high-frequency detail preservation, and a novel region transition smoothness loss that specifically addresses boundary artifacts between different lighting zones.
- Extensive evaluations on the LOL-v2, LSRW and our private datasets, demonstrating state-of-the-art performance. For instance, on our dataset, improvements of 3.26% in BRISQUE, 0.24% in NIQE, and 11.22% in PIQE were achieved over existing methods like Retinexformer and Zero-IG.

2. Related work

2.1. Retinex-based methods

The Retinex theory establishes a mathematical foundation for low-light image enhancement by decomposing images into two physical components: reflectance (representing intrinsic object properties) and illumination (encoding environmental lighting conditions). Its core hypothesis posits that adaptive image enhancement can be achieved through decoupling these components.

As the first practical implementation, Single-Scale Retinex (SSR) [JRW97] demonstrated the theory’s potential for global illumination adjustment, though its linear decomposition assumption often leads to detail loss. LIME [Guo16] pioneered illumination estimation as an independent optimization target, employing edge-preserving smoothing to generate initial illumination maps, which significantly improved local contrast. However, LIME tends to cause local over-exposure in high dynamic range scenarios. RUAS [LMZ*21] constructed a lightweight network with substantially reduced parameters through unrolled optimization and cooperative prior search, achieving real-time inference while maintaining performance. RetinexDIP [ZXW*21] reformulated the decomposition as a generative task using deep image priors to relax the strong coupling between reflectance and illumination. Nevertheless, its iterative optimization results in prolonged single-image processing time and sensitivity to initial illumination estimation errors. Retinexformer [CBL*23] introduced the Transformer architecture into the Retinex framework for the first time, modeling cross-region dependencies between illumination and reflectance via multi-head self-attention. Yet, its computational complexity grows quadratically with image resolution and requires extensive train-

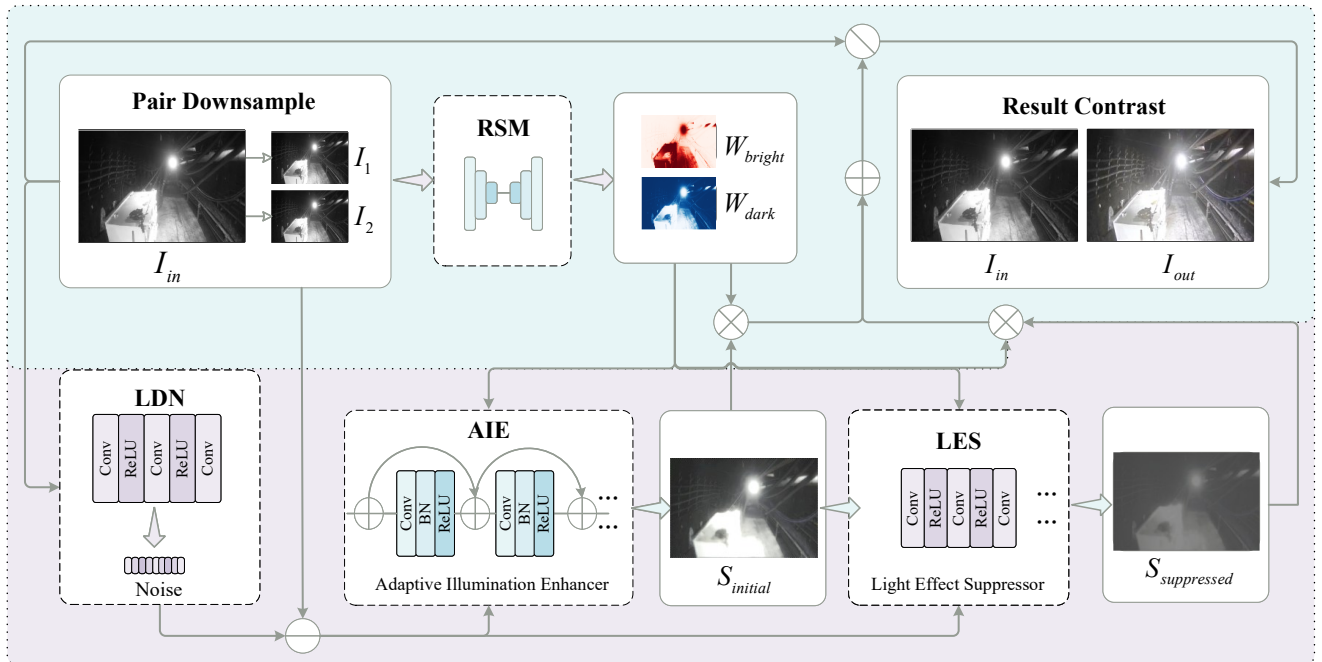


Figure 2: Overall architecture of the proposed region-adaptive low-light image enhancement framework, integrating region segmentation, adaptive illumination enhancement, and light effect suppression.

ing data. Diff-Retinex [YXZ*23] redefined the enhancement task through a conditional diffusion model, supplementing lost high-frequency information via generative diffusion after Retinex decomposition. ZERO-IG [SLZ*24] proposed an illumination-guided joint denoising and enhancement framework that enables end-to-end optimization without requiring any training data.

2.2. Learning-based models

In recent years, significant breakthroughs have been achieved in the field of low-light image enhancement, driven by advances in deep learning techniques. Algorithmic frameworks based on deep neural networks have substantially improved enhancement performance by learning complex nonlinear mappings from low-light to normal-light conditions in an end-to-end manner.

EnlightenGAN [JGL*21] overcomes the constraint of paired training data by integrating global-local discriminators, self-regularized perceptual loss, and attention mechanisms, significantly enhancing unsupervised enhancement quality. Zero-DCE [GLG*20] pioneered a zero-reference enhancement framework based on deep networks, optimizing dynamic range through adaptive curve adjustment; however, its performance is limited by reliance on multi-exposure training data, restricting generalization capability. SCI [MML*22] employs a weight-sharing self-calibrated illumination learning module to negotiate convergence between intermediate results at different stages, improving exposure stability while reducing computational overhead. The SKF framework [WPW*23] innovatively incorporates semantic segmentation priors to guide the enhancement process, effectively preserving structural details under low-light conditions. CLIP-LIT [LLZ*23] leverages the vision-language priors of the CLIP model, implementing a two-stage prompt learning strategy:

first generating initial prompt pairs to train the network, then iteratively refining the distribution gap between backlit, enhanced, and normal images for progressive performance improvement. UBF-Net [ZLH24] generates images with varying illumination and texture details through preprocessing, employs DenseNet for feature fusion to enhance contrast and brightness, and suppresses noise via a self-supervised denoising module based on U-Net++. The method proposed by Li et al [LXHL24] dynamically restores image dynamic range by estimating and correcting color shifts in overexposed and underexposed regions, thereby improving overall quality.

3. Method

3.1. Retinex theory foundation

Our method is grounded in Retinex theory [JRW97], which models human color constancy through the separation of reflectance and illumination components. According to this theory, any observed image can be decomposed into two multiplicative components:

$$\mathbf{I}_{in}(x, y) = \mathbf{R}(x, y) \odot \mathbf{S}(x, y) \quad (1)$$

where \mathbf{I}_{in} represents the input low-light image, \mathbf{R} denotes the reflectance component encoding intrinsic object properties, \mathbf{S} represents the illumination component capturing environmental lighting conditions, and \odot indicates element-wise multiplication.

In low-light scenarios, the illumination component \mathbf{S} contains very low values, causing the overall image to appear dark. The fundamental challenge in Retinex-based enhancement lies in accurately estimating \mathbf{S} from the single observed image \mathbf{I}_{in} , as this decomposition is inherently ill-posed—infinite combinations of \mathbf{R} and \mathbf{S} can produce the same observed image. Traditional Retinex methods

address this by imposing constraints such as spatial smoothness on illumination, but these approaches often fail in complex real-world scenarios with non-uniform illumination and strong light sources, leading to artifacts and detail loss.

Our approach leverages Retinex theory while addressing its limitations through region-adaptive processing. Instead of directly estimating the original illumination, we learn an enhanced illumination map \mathbf{S}_2 that compensates for poor lighting conditions while preserving the intrinsic reflectance properties:

$$\mathbf{I}_{\text{out}} = \mathbf{R} = \frac{\mathbf{I}_{\text{in}}}{\mathbf{S}_2} \quad (2)$$

where \mathbf{I}_{out} represents the final enhanced output image, \mathbf{R} denotes the recovered reflectance component (equivalent to the enhanced result), \mathbf{I}_{in} is the input low-light image, and \mathbf{S}_2 is our estimated enhanced illumination map. Here, \mathbf{S}_2 is estimated through our region-adaptive framework that distinguishes between dark regions requiring enhancement and bright regions needing artifact suppression. This approach enables more robust decomposition by incorporating spatial context and region-specific constraints.

3.2. Overall framework overview

Given a low-light input image $\mathbf{I}_{\text{in}} \in \mathbb{R}^{H \times W \times 3}$, our objective is to produce an enhanced image \mathbf{I}_{out} with superior visual quality. As illustrated in Figure 2, our framework integrates denoising processing, region segmentation, adaptive illumination enhancement, and light effect suppression modules, leveraging Retinex theory for image decomposition and end-to-end optimization. Algorithm 1 summarizes the step-by-step pipeline of our method.

The framework operates through four key stages. Initially, the input image \mathbf{I}_{in} undergoes preliminary processing by a lightweight denoising network LDN, yielding a reduced-noise image $\mathbf{L}_2 = \mathbf{I}_{\text{in}} - \text{LDN}(\mathbf{I}_{\text{in}})$. This step mitigates noise interference, providing cleaner input for subsequent illumination adjustments. Simultaneously, the region segmentation module analyzes \mathbf{I}_{in} to produce pixel-level weight maps: \mathbf{W}_{dark} for dark regions requiring enhancement and $\mathbf{W}_{\text{bright}}$ for bright regions needing overexposure suppression. Using \mathbf{L}_2 and the region weight maps, the Adaptive Illumination Enhancer estimates an initial illumination layer $\mathbf{S}_{\text{initial}}$, prioritizing brightness enhancement in dark regions. The Light Effect Suppressor then refines $\mathbf{S}_{\text{initial}}$ for bright regions, producing a suppressed illumination layer $\mathbf{S}_{\text{suppressed}}$ to mitigate halos and overexposure artifacts. These illumination components are fused using region weights to form the final illumination layer $\mathbf{S}_2 = \mathbf{S}_{\text{initial}} \odot \mathbf{W}_{\text{dark}} + \mathbf{S}_{\text{suppressed}} \odot \mathbf{W}_{\text{bright}}$, where \odot denotes element-wise multiplication.

Following Retinex theory, the final enhanced image is computed as $\mathbf{I}_{\text{out}} = \mathbf{I}_{\text{in}}/\mathbf{S}_2$ (or $\mathbf{L}_2/\mathbf{S}_2$), ensuring brightness enhancement in dark regions, natural suppression in bright regions, and preservation of structural details. The entire framework is trained in an unsupervised manner using a carefully designed composite loss function that integrates region-specific reconstruction, structural consistency, high-frequency detail preservation, and smoothness constraints to balance brightness enhancement, structural integrity, and artifact suppression.

Algorithm 1 Overall Framework

Require: Low-light input image $\mathbf{I}_{\text{in}} \in \mathbb{R}^{H \times W \times 3}$

Ensure: Enhanced output image \mathbf{I}_{out}

- 1: **Initial Denoising:** Apply lightweight denoising network LDN to obtain
 - 2: $\mathbf{L}_2 = \mathbf{I}_{\text{in}} - \text{LDN}(\mathbf{I}_{\text{in}})$
 - 3: **Region Segmentation:** Analyze \mathbf{I}_{in} to generate pixel-level region information
 - 4: Extract dark region weight map \mathbf{W}_{dark} and bright region weight map $\mathbf{W}_{\text{bright}}$
 - 5: **Adaptive Illumination Estimation:**
 - 6: Use AdaptiveEnhancer to estimate initial illumination layer:
 - 7: $\mathbf{S}_{\text{initial}} = \text{AdaptiveEnhancer}(\mathbf{L}_2, \mathbf{W}_{\text{dark}}, \mathbf{W}_{\text{bright}})$
 - 8: **Light Effect Suppression:**
 - 9: Apply LightEffectSuppressor to adjust bright regions:
 - 10: $\mathbf{S}_{\text{suppressed}} = \text{LightEffectSuppressor}(\mathbf{L}_2, \mathbf{S}_{\text{initial}}, \mathbf{W}_{\text{bright}})$
 - 11: **Illumination Fusion:** Combine enhanced and suppressed illumination maps
 - 12: $\mathbf{S}_2 = \mathbf{S}_{\text{initial}} \odot \mathbf{W}_{\text{dark}} + \mathbf{S}_{\text{suppressed}} \odot \mathbf{W}_{\text{bright}}$
 - 13: **Reflectance Recovery:** Apply Retinex decomposition to obtain final result
 - 14: $\mathbf{I}_{\text{out}} = \mathbf{I}_{\text{in}}/\mathbf{S}_2$ (or $\mathbf{L}_2/\mathbf{S}_2$)
 - 15: **return** \mathbf{I}_{out}
-

3.3. Region segmentation module

As shown in Figure 3, this module aims to divide the input image into primarily dark and primarily bright regions, and generate smooth membership weights for each pixel. The module operates through three main steps.

First, the input RGB image \mathbf{I}_{in} is converted into a grayscale intensity map \mathbf{I}_{gray} . We then set a threshold T_h (default is 0.5) and calculate the dark region weight \mathbf{W}_{dark} and bright region weight $\mathbf{W}_{\text{bright}}$ for each pixel using a Sigmoid function based on the deviation of its intensity value from T_h , to ensure a smooth transition of weights:

$$\mathbf{W}_{\text{dark}} = \sigma\left(\frac{T_h - \mathbf{I}_{\text{gray}}}{f_s}\right) \quad (3)$$

$$\mathbf{W}_{\text{bright}} = \sigma\left(\frac{\mathbf{I}_{\text{gray}} - T_h}{f_s}\right) \quad (4)$$

where $\sigma(\cdot)$ is the Sigmoid function, and f_s is a smoothing factor. To further avoid artifacts that may arise at region boundaries, we apply Gaussian smoothing to the generated weight maps. Finally, this module outputs the smooth weight maps \mathbf{W}_{dark} and $\mathbf{W}_{\text{bright}}$ for subsequent processing.

3.4. Adaptive illumination enhancer

As shown in Figure 4, the Adaptive Illumination Enhancer estimates and adaptively adjusts the illumination of the initially denoised image \mathbf{L}_2 based on region information. Algorithm 2 provides the implementation details.

The enhancer operates through a four-stage pipeline. First, base feature extraction is performed via a convolutional layer to obtain \mathbf{F}_{base} . Region-aware processing then separates \mathbf{F}_{base} into dark

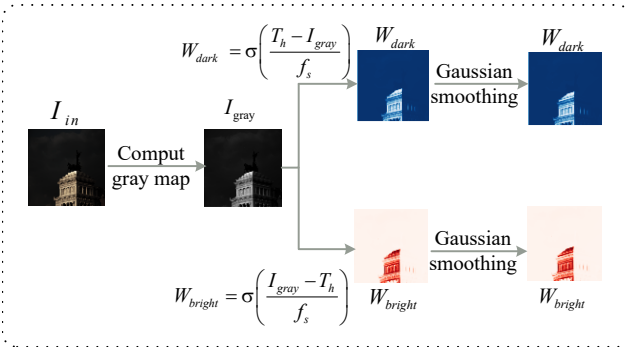


Figure 3: Region Segmentation Module (RSM).

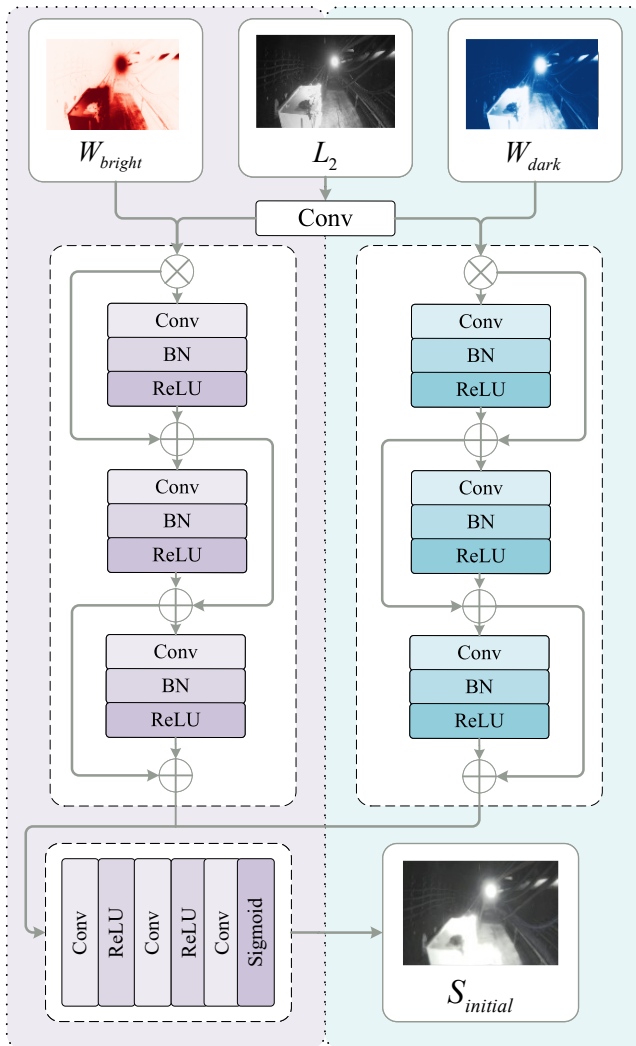


Figure 4: Adaptive Illumination Enhancer (AIE).

and bright region features (\mathbf{F}_{dark} and $\mathbf{F}_{\text{bright}}$) using weights from the region segmentation module. These features are independently enhanced through shared-weight residual convolutional blocks in parallel, enabling optimal feature transformations for different regions.

The processed features are subsequently concatenated and fed into an adaptive fusion module comprising convolutional layers to learn their optimal combination. Finally, an initial illumination map $\mathbf{S}_{\text{initial}}$ is generated via a convolutional layer followed by a Sigmoid activation function.

Algorithm 2 Adaptive Illumination Enhancer

Require: Denoised image L_2 , region weights \mathbf{W}_{dark} , $\mathbf{W}_{\text{bright}}$

Ensure: Initial illumination map $\mathbf{S}_{\text{initial}}$

- 1: **Base Feature Extraction:** Extract features using convolutional layer
- 2: $\mathbf{F}_{\text{base}} = \text{Conv}(L_2)$
- 3: **Region-Aware Processing:** Modulate features by region weights
- 4: $\mathbf{F}_{\text{dark}} = \mathbf{F}_{\text{base}} \odot \mathbf{W}_{\text{dark}}$
- 5: $\mathbf{F}_{\text{bright}} = \mathbf{F}_{\text{base}} \odot \mathbf{W}_{\text{bright}}$
- 6: **Parallel Region Enhancement:** Process features through shared-weight residual blocks
- 7: $\mathbf{F}_{\text{dark}}^{\text{enhanced}} = \text{ResidualBlocks}(\mathbf{F}_{\text{dark}})$
- 8: $\mathbf{F}_{\text{bright}}^{\text{enhanced}} = \text{ResidualBlocks}(\mathbf{F}_{\text{bright}})$
- 9: **Adaptive Fusion:** Concatenate and fuse processed features
- 10: $\mathbf{F}_{\text{fused}} = \text{AdaptiveFusion}(\text{Concat}(\mathbf{F}_{\text{dark}}^{\text{enhanced}}, \mathbf{F}_{\text{bright}}^{\text{enhanced}}))$
- 11: **Illumination Generation:** Generate initial illumination map
- 12: $\mathbf{S}_{\text{initial}} = \text{Sigmoid}(\text{Conv}(\mathbf{F}_{\text{fused}}))$
- 13: Clamp values to $[\epsilon, 1]$ where ϵ is a small positive constant
- 14: **return** $\mathbf{S}_{\text{initial}}$

3.5. Light effect suppressor

As shown in Figure 5, this module addresses over-bright areas and halo effects by specifically adjusting illumination estimation in bright regions. Algorithm 3 provides the implementation details.

The suppressor operates through a multi-stage process. Initially, the denoised image L_2 and the initial illumination estimate $\mathbf{S}_{\text{initial}}$ are concatenated along the channel dimension and fed into an encoder network composed of multiple convolutional layers with LeakyReLU activation functions to learn light effect patterns. The encoder outputs are then passed through an attention module (a convolutional layer followed by a Sigmoid function) to generate an attention map $\mathbf{A}_{\text{light}}$, which highlights potential light-effect regions. This attention mechanism uses convolutional operations to focus on spatial light effect patterns, differing from traditional self-attention approaches.

The attention map $\mathbf{A}_{\text{light}}$ serves a distinct purpose from $\mathbf{W}_{\text{bright}}$: while $\mathbf{W}_{\text{bright}}$ identifies potentially bright regions based on intensity thresholds, $\mathbf{A}_{\text{light}}$ learns to identify complex light effect patterns through feature analysis. The attention map is element-wise multiplied with the input bright region weights $\mathbf{W}_{\text{bright}}$ to produce combined weights $\mathbf{W}_{\text{combined}}$, ensuring that suppression applies only where both intensity-based and feature-based conditions are met.

The encoder features are then weighted by $\mathbf{W}_{\text{combined}}$ and fed into a suppressor network (convolutional layers with LeakyReLU activation followed by Sigmoid) to generate a pixel-wise suppression factor $\mathbf{F}_{\text{suppress}} \in [0, 1]$. Finally, the initial illumination map is adaptively adjusted via:

$$\mathbf{S}_{\text{suppressed}} = \mathbf{S}_{\text{initial}} \odot (1 - \mathbf{F}_{\text{suppress}} \odot \mathbf{W}_{\text{bright}}) \quad (5)$$

Through this framework, we obtain $\mathbf{S}_{\text{initial}}$ optimized for dark regions and $\mathbf{S}_{\text{suppressed}}$ optimized for bright regions, which are fused using region weights to form the final illumination map \mathbf{S}_2 that achieves balanced enhancement across different lighting conditions.

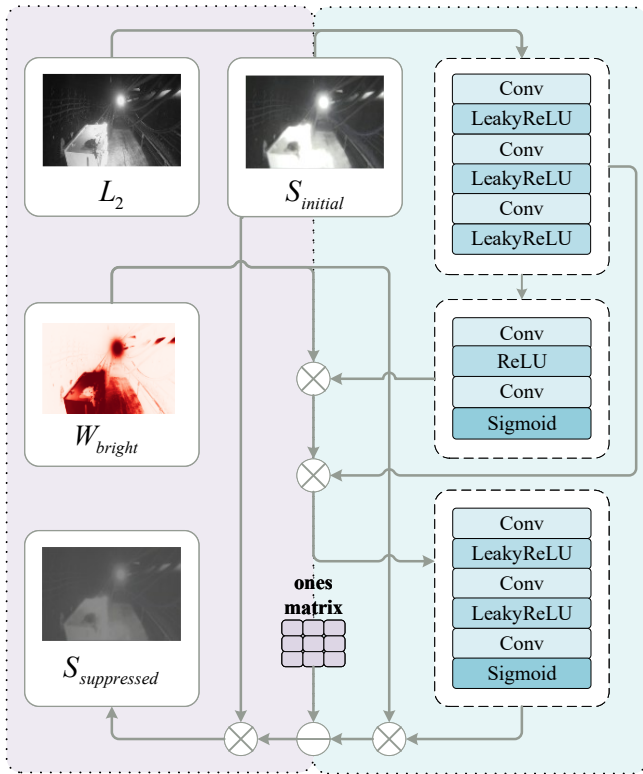


Figure 5: Light Effect Suppressor (LES).

3.6. Adaptive loss function

To effectively train the proposed network and balance different enhancement objectives, we design a novel adaptive loss function with multiple components that constrain the network's behavior from different dimensions. This multi-component design addresses key challenges: (1) simultaneously enhancing under-exposed regions while suppressing over-exposed areas; (2) handling extreme lighting conditions with point light sources; and (3) balancing brightness enhancement, artifact suppression, and detail preservation.

Region-adaptive reconstruction loss. The core idea of this loss term is to adaptively adjust the reconstruction target of the illumination map \mathbf{S}_2 based on regional information.

1. **Dark Region Enhancement:** For dark regions, we expect \mathbf{S}_2 to

Algorithm 3 Light Effect Suppressor

-
- Require:** Denoised image \mathbf{L}_2 , initial illumination $\mathbf{S}_{\text{initial}}$, bright weights $\mathbf{W}_{\text{bright}}$
- Ensure:** Suppressed illumination $\mathbf{S}_{\text{suppressed}}$
- 1: **Input Construction:** Concatenate input features
 - 2: $\mathbf{F}_{\text{input}} = \text{Concat}(\mathbf{L}_2, \mathbf{S}_{\text{initial}})$
 - 3: **Feature Encoding:** Learn light effect patterns using encoder network
 - 4: $\mathbf{F}_{\text{encoded}} = \text{Encoder}(\mathbf{F}_{\text{input}})$ {Multi-layer Conv + LeakyReLU}
 - 5: **Attention Guidance:** Generate attention map for light effect patterns
 - 6: $\mathbf{A}_{\text{light}} = \text{Sigmoid}(\text{Conv}(\mathbf{F}_{\text{encoded}}))$
 - 7: **Region-Weighted Attention:** Combine attention with bright region weights
 - 8: $\mathbf{W}_{\text{combined}} = \mathbf{A}_{\text{light}} \odot \mathbf{W}_{\text{bright}}$
 - 9: **Suppression Factor Generation:** Generate pixel-wise suppression factors
 - 10: $\mathbf{F}_{\text{weighted}} = \mathbf{F}_{\text{encoded}} \odot \mathbf{W}_{\text{combined}}$
 - 11: $\mathbf{F}_{\text{suppress}} = \text{Sigmoid}(\text{Suppressor}(\mathbf{F}_{\text{weighted}}))$ {Conv + LeakyReLU + Sigmoid}
 - 12: **Adaptive Illumination Suppression:** Apply suppression to initial illumination
 - 13: $\mathbf{S}_{\text{suppressed}} = \mathbf{S}_{\text{initial}} \odot (1 - \mathbf{F}_{\text{suppress}} \odot \mathbf{W}_{\text{bright}})$
 - 14: **return** $\mathbf{S}_{\text{suppressed}}$
-

significantly enhance the image brightness. The target brightness $\mathbf{L}_{\text{dark}}^*$ is computed dynamically as:

$$\mathbf{L}_{\text{dark}}^* = \mathbf{L}_2 \cdot \alpha \cdot \mathbf{W}_{\text{dark}} \quad (6)$$

where $\alpha = 0.5/(\mu_Y + \epsilon)$ is the enhancement factor computed from the average luminance μ_Y of each input image, and $\epsilon = 1e^{-9}$ prevents division by zero. The loss is:

$$\mathcal{L}_{\text{dark_rec}} = \|\mathbf{S}_2 \odot \mathbf{W}_{\text{dark}} - (\mathbf{L}_{\text{dark}}^* \odot \mathbf{W}_{\text{dark}})\|_2^2 \quad (7)$$

2. **Bright Region Preservation:** For bright regions, we prefer to preserve the original brightness rather than over-enhancing. The goal is for \mathbf{S}_2 in bright regions to approximate the illumination level of the denoised image \mathbf{L}_2 :

$$\mathcal{L}_{\text{bright_rec}} = \|\mathbf{S}_2 \odot \mathbf{W}_{\text{bright}} - (\mathbf{L}_2 \odot \mathbf{W}_{\text{bright}})\|_2^2 \quad (8)$$

Structural and high-frequency detail preservation loss. To ensure that the structural information and high-frequency details of the enhanced image are preserved, we introduce:

1. **Structural Consistency Loss:** This loss adopts SSIM-like computation between the input image \mathbf{I}_{in} and enhanced output \mathbf{I}_{out} :

$$\text{SSIM}(\mathbf{I}_{\text{in}}, \mathbf{I}_{\text{out}}) = \frac{(2\mu_1\mu_2 + C_1)(2\sigma_{12} + C_2)}{(\mu_1^2 + \mu_2^2 + C_1)(\sigma_1^2 + \sigma_2^2 + C_2)} \quad (9)$$

where μ_1, μ_2 are local means, σ_1^2, σ_2^2 are local variances, σ_{12} is local covariance, and C_1, C_2 are stability constants. The region-weighted structural loss is:

$$\mathcal{L}_{\text{struct}} = \mathbb{E}[(1 - \text{SSIM}(\mathbf{I}_{\text{in}}, \mathbf{I}_{\text{out}})) \odot (\mathbf{W}_{\text{dark}} + \lambda_s \mathbf{W}_{\text{bright}})] \quad (10)$$

where λ_s is a balancing factor.

2. **High-Frequency Consistency Loss:** This loss extracts high-frequency components of the input image \mathbf{I}_{in} and the enhanced output \mathbf{I}_{out} using a Laplacian operator ∇^2 and minimizes the L1 difference between them. Region weights are used with emphasis on preserving details in dark regions:

$$\mathcal{L}_{hf} = \mathbb{E} \left[\left| \nabla^2 \mathbf{I}_{in} - \nabla^2 \mathbf{I}_{out} \right| \odot (\lambda_{hf_d} \mathbf{W}_{dark} + \lambda_{hf_b} \mathbf{W}_{bright}) \right] \quad (11)$$

where λ_{hf_d} and λ_{hf_b} are weighting factors for dark and bright regions, respectively.

Smoothness and regularization loss. To maintain visual coherence and avoid artifacts in the enhanced illumination map \mathbf{S}_2 , we introduce:

1. **Illumination Smoothness Loss:** We follow common smoothness losses [LB14] and Total Variation Loss (\mathcal{L}_{TV_loss}) [OBG*05] to constrain the smoothness of the generated illumination map \mathbf{S}_2 , avoiding abrupt changes in brightness.

$$\begin{aligned} \mathcal{L}_{S_smooth} &= \text{SmoothLoss}(\mathbf{L}_2, \mathbf{S}_2) \\ \mathcal{L}_{S_tv} &= \mathcal{L}_{TV}(\mathbf{S}_2) \end{aligned} \quad (12)$$

2. **Region Transition Smoothness Loss:** This is a novel loss term specifically designed to ensure that the illumination map \mathbf{S}_2 remains smooth in the transition zones between dark and bright regions. It first identifies region boundaries (by comparing the differences between \mathbf{W}_{dark} and \mathbf{W}_{bright}), then calculates the gradient of \mathbf{S}_2 in these boundary regions and minimizes these gradient values:

$$\mathcal{L}_{trans_smooth} = \mathbb{E}_{\text{boundary}} [|\nabla_x \mathbf{S}_2| + |\nabla_y \mathbf{S}_2|] \quad (13)$$

Initial denoising consistency loss. To ensure the effectiveness of the initial denoising network LDN, its training borrows the loss function concept from the LD-Net in Zero-IG [SLZ*24]. The input low-light image \mathbf{I}_{in} is first decomposed into two independent low-resolution images, $L_{11} = G_1(\mathbf{I}_{in})$ and $L_{12} = G_2(\mathbf{I}_{in})$, via a downsampling operation $G = (G_1, G_2)$. The network $\text{LDN}(\cdot)$ aims to predict the noise in the input image. Its loss function, $\mathcal{L}_{denoise}$, consists of the following two parts:

1. **Residual Loss:** This loss encourages the network to learn noise that can explain the difference between the two downsampled images.

$$\mathcal{L}_{LD_res} = \|L_{11} - \text{LDN}(L_{11}) - L_{12}\|_2^2 + \|L_{12} - \text{LDN}(L_{12}) - L_{11}\|_2^2 \quad (14)$$

where $\text{LDN}(L_{11})$ and $\text{LDN}(L_{12})$ are the noise predicted by the network for the downsampled images L_{11} and L_{12} , respectively.

2. **Consistency Loss:** This loss ensures consistency in denoising across different scales.

$$\begin{aligned} \mathcal{L}_{LD_cons} &= \|(L_{11} - \text{LDN}(L_{11})) - G_1(\mathbf{I}_{in} - \text{LDN}(\mathbf{I}_{in}))\|_2^2 \\ &\quad + \|(L_{12} - \text{LDN}(L_{12})) - G_2(\mathbf{I}_{in} - \text{LDN}(\mathbf{I}_{in}))\|_2^2 \end{aligned} \quad (15)$$

where $\text{LDN}(\mathbf{I}_{in})$ is the noise predicted by the network for the original input image \mathbf{I}_{in} .

The total loss for the initial denoising network is:

$$\mathcal{L}_{denoise} = \mathcal{L}_{LD_res} + \mathcal{L}_{LD_cons} \quad (16)$$

By minimizing this loss, the LDN network learns to remove noise, outputting an image $\mathbf{L}_2 = \mathbf{I}_{in} - \text{LDN}(\mathbf{I}_{in})$ with a lower noise level for subsequent processing.

The final total loss combines all components:

$$\begin{aligned} \mathcal{L}_{total} &= \lambda_1 \mathcal{L}_{dark_rec} + \lambda_2 \mathcal{L}_{bright_rec} + \lambda_3 \mathcal{L}_{struct} + \lambda_4 \mathcal{L}_{hf} \\ &\quad + \lambda_5 \mathcal{L}_{S_smooth} + \lambda_6 \mathcal{L}_{S_tv} \\ &\quad + \lambda_7 \mathcal{L}_{trans_smooth} + \lambda_8 \mathcal{L}_{LD_res} + \lambda_9 \mathcal{L}_{LD_cons} \end{aligned} \quad (17)$$

where the specific loss weights employed are: $\lambda_1 = 700$, $\lambda_2 = 500$, $\lambda_3 = 200$, $\lambda_4 = 100$, $\lambda_5 = 5$, $\lambda_6 = 1600$, $\lambda_7 = 50$, $\lambda_8 = \lambda_9 = 1000$. These weights were determined through systematic experimentation to achieve optimal balance across all enhancement objectives.

4. Experiments

4.1. Experimental setup

Benchmarks. To validate the generalization capability of our proposed method, we conduct experiments on two widely used public datasets (LOL-v2 [WWYL18] and LSRW [HXY*23]) and a private dataset from real-world underground coal mine environments.

LOL-v2 comprises LOL-v2-real (789 pairs) and LOL-v2-synthetic (1000 pairs) subsets, while LSRW includes LSRW-Huawei (2480 pairs) and LSRW-Nikon (3170 pairs) variants.

Private Underground Coal Mine Dataset: We collected 3500 low-light images (600×400) from a real underground coal mine surveillance system. These images present unique challenges with strong point light sources, uneven illumination, and complex industrial scenes, forming a demanding testbed that significantly differs from conventional datasets.

Parameter settings. In the experimental environment setting, we use the NVIDIA RTX 4090 GPU as the computing platform and implement the algorithm within the PyTorch 1.13.1 framework. Additionally, the optimizer is configured as Adam with a learning rate of 0.0002, the batch size is set to 1, and the number of training epochs is defined as 2000.

Baselines. To comprehensively evaluate the performance of the proposed method, this study selects six representative low-light image enhancement approaches as baseline comparisons, including six learning methods (RUAS [LMZ*21], SCI [MML*22], ZeroDCE++ [GLG*20], EnlightenGAN [JGL*21], NeRCo [YDW*23], and Zero-IG [SLZ*24]). All comparative experiments are implemented using the official open-source codes of these methods, strictly following the parameter configurations recommended in their original publications. The experiments are conducted under a unified hardware environment to ensure the fairness and reproducibility of the experimental results.

Metrics. To comprehensively evaluate the enhancement results, we employ three objective no-reference image quality assessment metrics: BRISQUE [MMB12], NIQE [MSB12], and PIQE [VPB*15]. Our choice of these metrics over traditional reference-based metrics such as PSNR or SSIM is well-motivated. First, our private coal mine dataset lacks ground truth normal-light references, necessitating no-reference metrics that assess image quality without paired

Table 1: Quantitative comparison of image quality metrics on multiple datasets (Ours-Dataset, LOL-v2-Real, LOL-v2-Syn, LSRW-Huawei, LSRW-Nikon). The best unsupervised results are in **bold**, and the second-best are underlined.

Datasets	Metrics	RUAS cvpr2021	SCI cvpr2022	ZeroDCE++ TRAMI2021	EnlightenGAN TIP2022	NeRCo ICCV2023	Zero-IG-d cvpr2024	Our
Ours-Dataset	<i>brisque</i>	26.59	25.44	<u>18.38</u>	19.94	22.63	32.06	17.78
	<i>niqe</i>	4.86	4.96	4.41	4.68	<u>4.23</u>	5.07	4.22
	<i>piqe</i>	28.22	26.48	<u>25.24</u>	30.45	36.32	37.51	22.41
LOL-v2-Real	<i>brisque</i>	20.85	32.51	38.35	17.42	18.68	28.15	<u>18.64</u>
	<i>niqe</i>	7.47	8.05	8.18	7.12	<u>7.11</u>	7.96	7.09
	<i>piqe</i>	24.62	40.00	47.48	28.15	29.42	<u>23.42</u>	23.17
LOL-v2-Syn	<i>brisque</i>	38.60	22.81	22.89	18.62	<u>18.54</u>	38.04	18.14
	<i>niqe</i>	6.51	4.87	4.83	<u>4.49</u>	4.50	5.08	4.47
	<i>piqe</i>	50.04	46.49	43.67	40.42	36.35	48.12	<u>40.02</u>
LSRW-Huawei	<i>brisque</i>	42.39	24.36	27.69	21.42	<u>21.36</u>	34.76	21.04
	<i>niqe</i>	5.60	4.25	4.65	<u>4.13</u>	4.22	4.85	4.07
	<i>piqe</i>	55.55	33.58	25.64	27.58	39.93	<u>27.55</u>	32.07
LSRW-Nikon	<i>brisque</i>	45.37	28.70	30.27	24.94	<u>24.91</u>	35.39	21.11
	<i>niqe</i>	5.13	3.94	3.61	4.53	<u>3.73</u>	4.18	4.26
	<i>piqe</i>	59.75	33.84	34.12	33.94	47.89	<u>32.66</u>	32.17

data. Second, our unsupervised training framework focuses on perceptual quality rather than pixel-level accuracy, making perceptual quality metrics more appropriate for evaluation. Notably, all three metrics exhibit negative correlation characteristics, meaning lower scores correspond to better image quality.

4.2. Benchmark evaluations

Visual comparisons. Figure 7 and Figure 8 present qualitative comparisons on representative samples from the LOL-v2 and LSRW datasets. Existing methods struggle to balance brightness enhancement with detail preservation across varying lighting conditions. In contrast, our method achieves a more balanced enhancement between global brightness improvement and artifact suppression, effectively preserving texture details in both bright and dark regions while maintaining natural visual quality. For example, on images from the LOL-v2 dataset, the algorithm enhances the visibility of dark-region walls while improving the readability of books in bright regions—no overexposure is observed, the text textures are well-preserved, and no color distortion occurs. These results demonstrate the generalization capability of our algorithm in complex lighting scenarios.

Furthermore, Figure 1 and Figure 9 show qualitative comparisons on representative samples from the underground coal mine low-light dataset constructed in this study. The input images, captured in real mining environments, are characterized by strong point light sources and uneven illumination. Existing methods face challenges in this scenario: RUAS, SCI, and ZeroDCE++ produce prominent halo effects around strong light sources and over-enhance bright regions; EnlightenGAN generates unnatural color enhancements; NeRCo and Zero-IG fail to effectively suppress diffusion artifacts from strong light sources, causing unnatural distortions in high-brightness areas. Our method, however, successfully suppresses halo effects, enhances visibility in dark regions, and maintains an overall natural visual quality, demonstrating superior performance.

Table 2: Runtime, Parameters, and FLOPs efficiency comparison of different methods on private coal mine dataset.

Methods	Runtime (ms)	Parameters (M)	FLOPs (G)
RUAS	16.2	0.004	0.89
SCI	-	-	-
ZeroDCE++	18.4	0.167	2.05
EnlightenGAN	48.3	7.921	45.32
NeRCo	35.2	2.843	12.86
Zero-IG	31.5	0.327	4.82
Our	<u>17.8</u>	<u>0.156</u>	<u>1.95</u>

Quantitative comparisons. As summarized in Table 1, our method achieves consistently superior performance across most datasets and metrics.

On our private dataset, our method demonstrates excellence across all three metrics, achieving improvements of 3.26% in BRISQUE, 0.24% in NIQE, and 11.22% in PIQE compared to the second-best methods. These improvements highlight our method's effectiveness in handling complex real-world scenarios with strong point light sources and uneven illumination.

On public benchmarks, our method similarly demonstrates competitive advantages. On the LOL-v2-Real dataset, our method achieves the best performance in both NIQE and PIQE metrics, with improvements of 0.28% and 1.07%, respectively, over the second-best methods, and attains the second-best result in BRISQUE. On the LOL-v2-Syn dataset, our method achieves the best performance in BRISQUE and NIQE metrics, improving by 2.16% and 0.45% compared to the second-best methods, and ranks second in PIQE. On the LSRW-Huawei dataset, our method achieve the best results in BRISQUE and NIQE metrics, with improvements of 1.50% and 1.45%, respectively. on the LSRW-Nikon dataset, our method achieves the best performance in BRISQUE and PIQE metrics, with improvements of 15.25% and 1.50%.

Regarding computational efficiency, Table 2 shows our method achieves competitive performance with 17.8ms runtime, 0.156M parameters, and 1.95G FLOPs. Although RUAS demonstrates the highest efficiency with 16.2ms runtime, 0.004M parameters, and 0.89G FLOPs, it significantly underperforms in image quality metrics as demonstrated in Table 1. Moreover, while SCI shows lower computational complexity, training collapse is commonly observed during its training process. Therefore, for the sake of fairness, we have excluded it from the comparison. Our approach strikes an optimal balance between enhancement quality and computational efficiency, adding only 1.6ms compared to RUAS while delivering substantial quality improvements. This efficiency profile makes our framework well-suited for real-time deployment in safety-critical applications.

Overall, the quantitative evaluation confirms our approach's superiority across diverse lighting scenarios, demonstrating its effectiveness for low-light image enhancement in complex real-world applications.

4.3. Ablation study

Subnetworks ablation study. To evaluate the importance of each subnetwork, we progressively added the Adaptive Illumination Enhancer (AIE), Region Segmentation Module (RSM), Lightweight Denoising Network (LDN), and Light Effect Suppressor (LES). Table 3 presents quantitative results on both private and public datasets, while Figure 6a shows qualitative comparisons.

We first evaluate the effect of progressively adding subnetworks on the private dataset. As shown in Table 3, using only AIE yields a BRISQUE score of 25.50, NIQE of 4.43, and PIQE of 39.12, indicating poor image quality. Incorporating RSM improves BRISQUE to 20.97 and PIQE to 32.37, enhancing image quality, though NIQE slightly increases to 4.65, suggesting that RSM aids in distinguishing dark and bright regions. Adding LDN slightly reduces BRISQUE to 20.71, NIQE to 4.61, and PIQE to 32.33, demonstrating a denoising effect. When all modules are integrated, BRISQUE drops significantly to 17.78, NIQE to 4.22, and PIQE to 22.41, marking a substantial improvement in image quality. As illustrated in Figure 6a, AIE alone results in severe color deviation from the original image with significant distortion; adding RSM leads to over-enhancement and detail loss in bright regions; incorporating LDN reduces noise but over-enhancement in bright regions persists; while the complete model effectively enhances dark regions, suppresses over-enhancement in bright areas, and preserves details and textures.

On the public dataset, the ablation study similarly validates the role of each subnetwork. Table 3 shows that using AIE alone results in a BRISQUE of 22.94, NIQE of 4.52, and PIQE of 41.92, reflecting suboptimal image quality. Adding RSM improves BRISQUE to 20.55 and PIQE to 41.80, indicating enhanced image quality, though NIQE slightly increases to 4.56. Incorporating LDN further reduces BRISQUE to 20.42, NIQE to 4.54, and PIQE to 40.70, evidencing a denoising effect. With all modules integrated, BRISQUE significantly improves to 18.14, NIQE to 4.47, and PIQE to 40.02, demonstrating substantial enhancement. These findings underscore the critical role of LES in suppressing over-enhancement in bright regions and preserving details, while LDN effectively reduces noise.

Table 3: Ablation study on subnetworks. Quantitative comparison of image quality metrics with different module combinations on private and public datasets.

Subnetworks				Private			Public		
AIE	RSM	LDN	LES	brisque	niqe	piqe	brisque	niqe	piqe
✓	×	×	×	25.50	4.43	39.12	22.94	4.52	41.92
✓	✓	×	×	20.97	4.65	32.37	20.55	4.56	41.80
✓	✓	✓	×	20.71	4.61	32.33	20.42	4.54	40.70
✓	✓	✓	✓	17.78	4.22	22.41	18.14	4.47	40.02

Loss functions ablation study. We assessed the importance of each loss component by progressively adding Region-Adaptive Reconstruction Loss (RARL), Smoothness and Regularization Loss (SRL), Structural and High-Frequency Detail Preservation Loss (SHDPL), and Initial Denoising Consistency Loss (IDCL). Results are shown in Table 4 and Figure 6b.

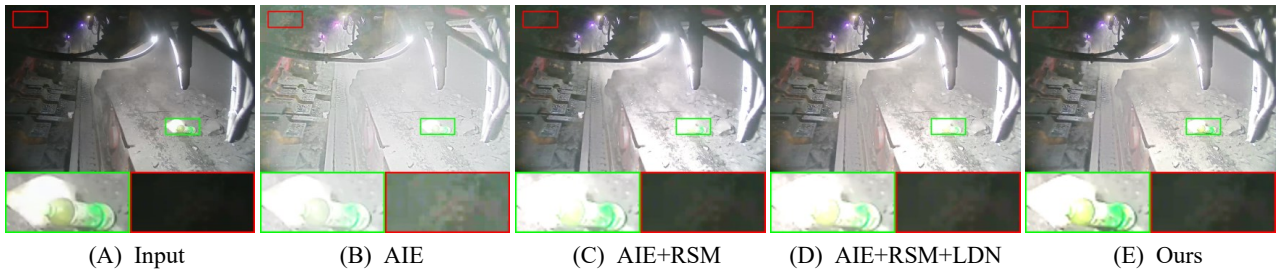
On the private dataset, using only RARL yields poor quality (BRISQUE: 27.32, NIQE: 4.62, PIQE: 38.72). Adding SRL improves metrics to 26.93, 4.56, and 35.82 respectively. Incorporating SHDPL significantly enhances quality (19.83, 4.31, 26.93), while the complete loss function achieves optimal results (17.78, 4.22, 22.41). As demonstrated in Figure 6b, using RARL alone results in severe overexposure, loss of detail and texture, and significant distortion; adding SRL reduces distortion with partial recovery of texture details, though overexposure persists; incorporating SHDPL restores texture details and suppresses overexposure; while the complete loss function not only resolves texture detail and overexposure issues but also effectively reduces noise.

On the public dataset, using RARL alone results in a BRISQUE of 63.95, NIQE of 9.88, and PIQE of 67.92, indicating extremely poor image quality. Adding SRL improves BRISQUE to 52.54, NIQE to 8.52, and PIQE to 63.47, enhancing quality. Incorporating SHDPL significantly reduces BRISQUE to 20.21, NIQE to 4.63, and PIQE to 41.65, showing further improvement. With all losses integrated, BRISQUE drops to 18.14, NIQE to 4.47, and PIQE to 40.02, marking a substantial quality boost. These results highlight SHDPL's critical role in preserving details and suppressing overexposure, with IDCL contributing significantly to noise reduction.

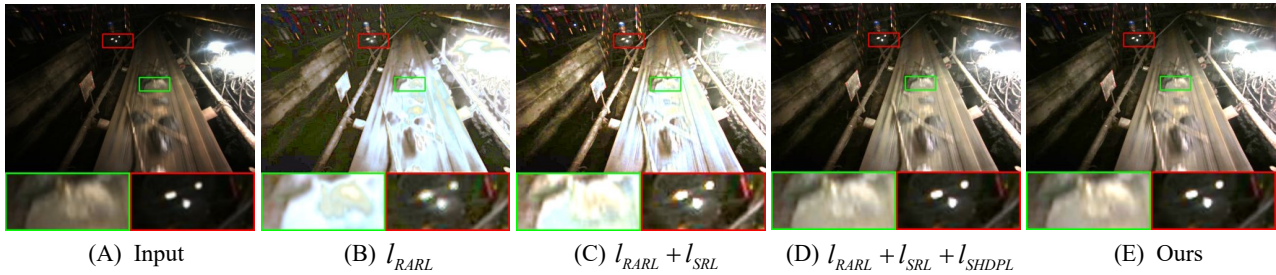
Table 4: Ablation study on loss functions. Quantitative comparison of image quality metrics with different loss components on private and public datasets.

Loss Functions				Private			Public		
RARL	SRL	SHDPL	IDCL	brisque	niqe	piqe	brisque	niqe	piqe
✓	×	×	×	27.32	4.62	38.72	63.95	9.88	67.92
✓	✓	×	×	26.93	4.56	35.82	52.54	8.52	63.47
✓	✓	✓	×	19.83	4.31	26.93	20.21	4.63	41.65
✓	✓	✓	✓	17.78	4.22	22.41	18.14	4.47	40.02

The ablation studies confirm that each subnetwork and loss function component in our proposed framework is essential for achieving high-quality low-light image enhancement. Notably, LES and SHDPL are pivotal in suppressing over-enhancement in bright regions and preserving details. This synergy ensures effective en-



(a) Ablation study on subnetworks. Visual comparison of results with different module combinations, evaluated on private datasets.



(b) Ablation study on loss functions. Visual comparison of results with different loss components, evaluated on private datasets.

Figure 6: Ablation studies comparing subnetwork module combinations and loss function components, evaluated on private datasets.

hancement of dark regions, natural suppression of bright areas, and retention of overall image details.

5. Conclusion

We propose a novel region-adaptive low-light image enhancement framework that effectively addresses unnatural artifacts, overexposure, and detail loss under complex lighting conditions. By integrating a Retinex-based decomposition approach with region-specific processing (including a Region Segmentation Module, Adaptive Illumination Enhancer, Light Effect Suppressor, and Composite Loss Function), our method achieves a balanced enhancement of brightness, artifact suppression, and texture preservation. Extensive experiments on the LOL-v2, LSRW, and our underground coal mine dataset demonstrate superior visual quality. The method shows improvements in BRISQUE, NIQE, and PIQE metrics compared to existing approaches. Ablation studies confirm the critical role of each component in achieving these results.

Acknowledgements

This project was partially supported by the National Natural Science Foundation of China under Grant 62473125, 62033012, the Ecological Environment Research Project of Anhui Province under Grant 2023hb0006, and the Anhui Provincial Key Research and Development Program (No. 2022107020030).

References

[CBL*23] CAI, YUANHAO, BIAN, HAO, LIN, JING, et al. “Retinexformer: One-stage retinex-based transformer for low-light image enhancement”. *Proceedings of the IEEE/CVF international conference on computer vision*. 2023, 12504–12513 2.

[CGZ18] CAI, JIANRUI, GU, SHUHANG, and ZHANG, LEI. “Learning a deep single image contrast enhancer from multi-exposure images”. *IEEE Transactions on Image Processing* 27.4 (2018), 2049–2062 1.

[CS04] CHENG, HENG-DA and SHI, XJ. “A simple and effective histogram equalization approach to image enhancement”. *Digital signal processing* 14.2 (2004), 158–170 1.

[GLG*20] GUO, CHUNLE, LI, CHONGYI, GUO, JICHANG, et al. “Zero-reference deep curve estimation for low-light image enhancement”. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020, 1780–1789 1, 3, 7.

[Guo16] GUO, XIAOJIE. “LIME: A method for low-light image enhancement”. *Proceedings of the 24th ACM international conference on Multimedia*. 2016, 87–91 1, 2.

[HCC12] HUANG, SHIH-CHIA, CHENG, FAN-CHIEH, and CHIU, YI-SHENG. “Efficient contrast enhancement using adaptive gamma correction with weighting distribution”. *IEEE transactions on image processing* 22.3 (2012), 1032–1041 1.

[HXY*23] HAI, JIANG, XUAN, ZHU, YANG, REN, et al. “R2rnet: Low-light image enhancement via real-low to real-normal network”. *Journal of Visual Communication and Image Representation* 90 (2023), 103712 7.

[JB25] JOSHUA, A SAM and BALASUBRAMANIAM, P. “An adaptive low-light image enhancement method via fusion of a new intuitionistic fuzzy generator and fractal-fractional derivative”. *Signal, Image and Video Processing* 19.3 (2025), 233 1.

[JGL*21] JIANG, YIFAN, GONG, XINYU, LIU, DING, et al. “Enlightengan: Deep light enhancement without paired supervision”. *IEEE transactions on image processing* 30 (2021), 2340–2349 1, 3, 7.

[JRW97] JOHNSON, DANIEL J, RAHMAN, ZIA-UR, and WOODDELL, GLENN A. “Properties and performance of a center/surround retinex”. *IEEE transactions on image processing* 6.3 (1997), 451–462 2, 3.

[LB14] LI, YU and BROWN, MICHAEL S. “Single image layer separation using relative smoothness”. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014, 2752–2759 7.

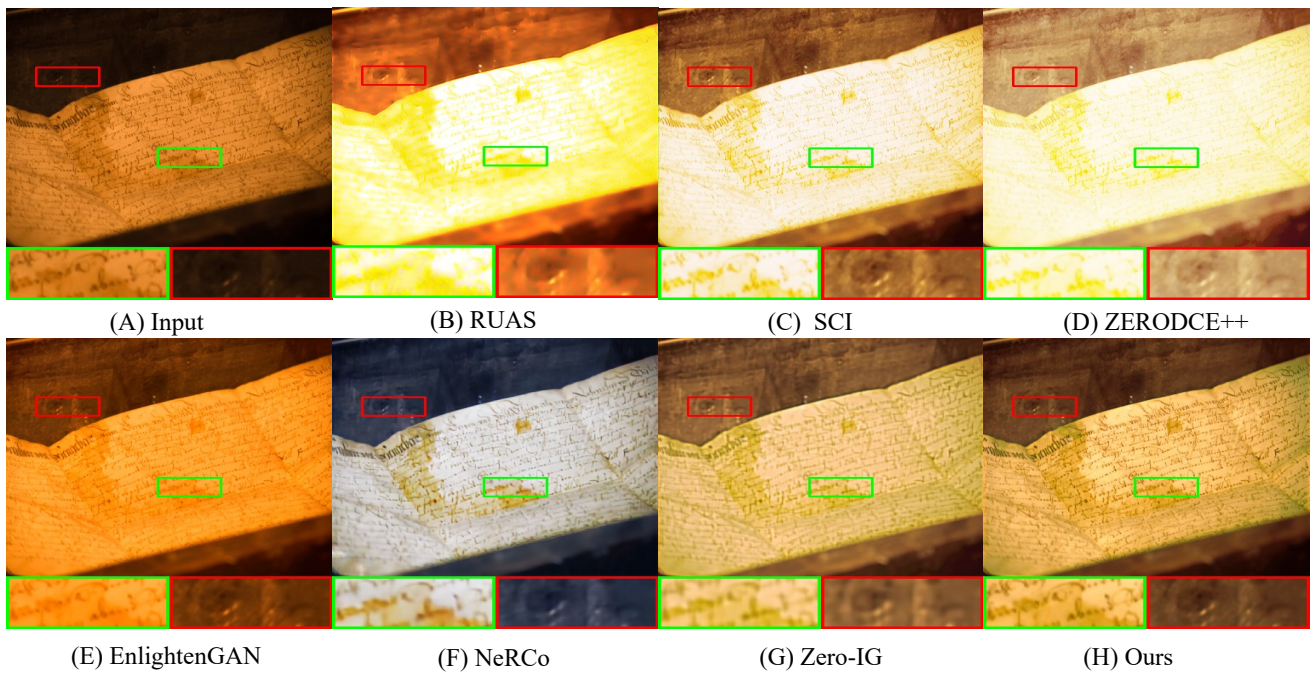


Figure 7: Visual comparison on the LOL-v2 dataset.

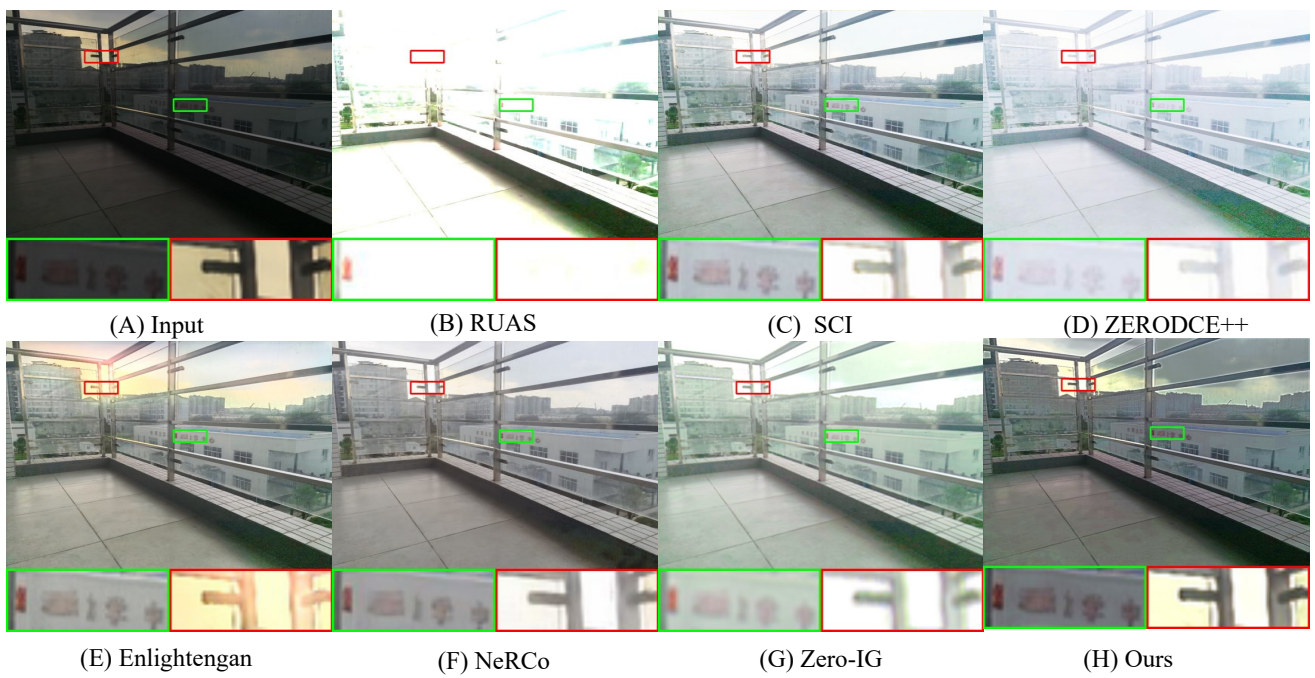


Figure 8: Visual comparison on the LSRW dataset.

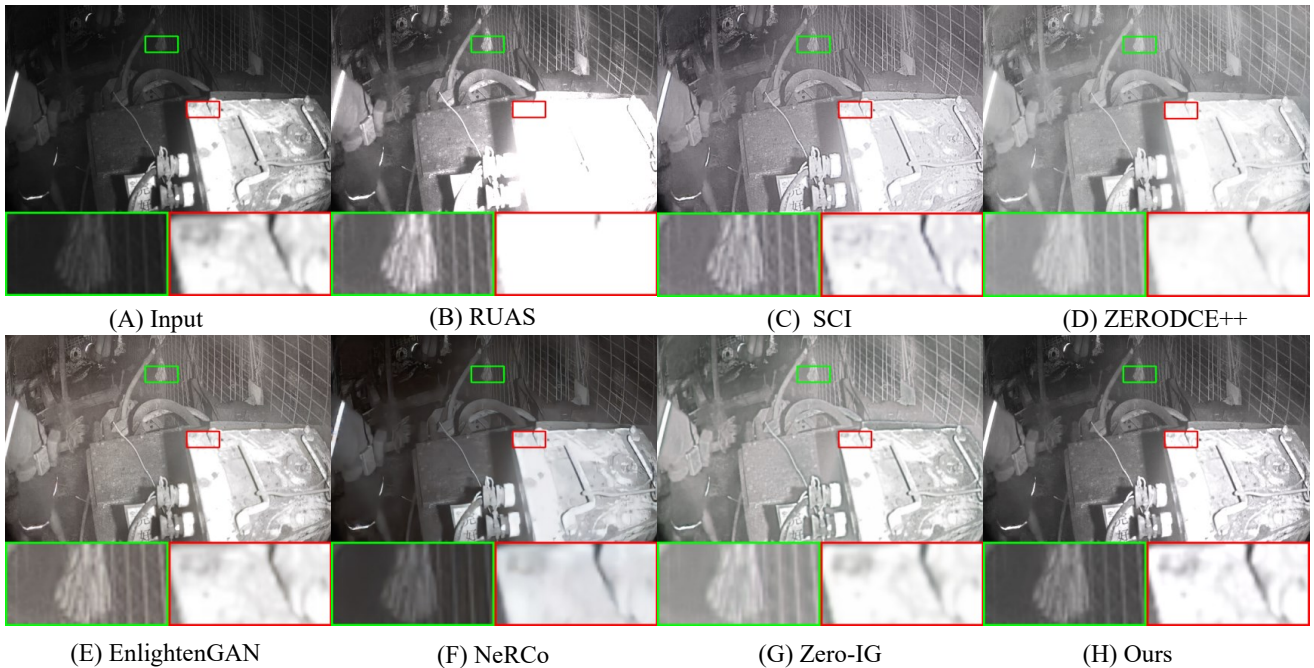


Figure 9: Visual comparison on the underground coal mine dataset.

[LLZ*23] LIANG, ZHEXIN, LI, CHONGYI, ZHOU, SHANGCHEN, et al. “Iterative prompt learning for unsupervised backlit image enhancement”. *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2023, 8094–8103 3.

[LMZ*21] LIU, RISHENG, MA, LONG, ZHANG, JIAAO, et al. “Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement”. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2021, 10561–10570 1, 2, 7.

[LSLH13] LEE, CHANG-HSING, SHIH, JAU-LING, LIEN, CHENG-CHANG, and HAN, CHIN-CHUAN. “Adaptive multiscale retinex for image contrast enhancement”. *2013 International Conference on Signal-Image Technology & Internet-Based Systems*. IEEE. 2013, 43–50 1.

[LXHL24] LI, YIYU, XU, KE, HANCKE, GERHARD PETRUS, and LAU, RYNSON WH. “Color shift estimation-and-correction for image enhancement”. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024, 25389–25398 3.

[MMB12] MITTAL, ANISH, MOORTHY, ANUSH KRISHNA, and BOVIK, ALAN CONRAD. “No-reference image quality assessment in the spatial domain”. *IEEE Transactions on image processing* 21.12 (2012), 4695–4708 7.

[MML*22] MA, LONG, MA, TENGJU, LIU, RISHENG, et al. “Toward fast, flexible, and robust low-light image enhancement”. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022, 5637–5646 1, 3, 7.

[MSB12] MITTAL, ANISH, SOUNDARARAJAN, RAJIV, and BOVIK, ALAN C. “Making a “completely blind” image quality analyzer”. *IEEE Signal processing letters* 20.3 (2012), 209–212 7.

[OBG*05] OSHER, STANLEY, BURGER, MARTIN, GOLDFARB, DONALD, et al. “An iterative regularization method for total variation-based image restoration”. *Multiscale Modeling & Simulation* 4.2 (2005), 460–489 7.

[SLZ*24] SHI, YIQI, LIU, DUO, ZHANG, LIGUO, et al. “Zero-ig: zero-shot illumination-guided joint denoising and adaptive enhancement for low-light images”. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2024, 3015–3024 3, 7.

[SY24] SU, YONGGANG and YANG, XUEJIE. “Low-light image enhancement based on variational image decomposition”. *Multimedia Systems* 30.6 (2024), 383 1.

[VPB*15] VENKATANATH, NARASIMHAN, PRANEETH, D, BH, MARUTHI CHANDRASEKHAR, et al. “Blind image quality evaluation using perception based features”. *2015 twenty first national conference on communications (NCC)*. IEEE. 2015, 1–6 7.

[WPW*23] WU, YUHUI, PAN, CHEN, WANG, GUOQING, et al. “Learning semantic-aware knowledge guidance for low-light image enhancement”. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023, 1662–1671 3.

[WWYL18] WEI, CHEN, WANG, WENJING, YANG, WENHAN, and LIU, JIAYING. “Deep retinex decomposition for low-light enhancement”. *arXiv preprint arXiv:1808.04560* (2018) 7.

[XHZ*24] XU, LINTAO, HU, CHANGHUI, ZHU, WEIHONG, et al. “Degrade for upgrade: Learning degradation representations for real-world low-light image enhancement”. *Computers and Electrical Engineering* 119 (2024), 109622 1.

[YDW*23] YANG, SHUZHOU, DING, MOXUAN, WU, YANMIN, et al. “Implicit neural representation for cooperative low-light image enhancement”. *Proceedings of the IEEE/CVF international conference on computer vision*. 2023, 12918–12927 7.

[YXZ*23] YI, XUNPENG, XU, HAN, ZHANG, HAO, et al. “Diff-retinex: Rethinking low-light image enhancement with a generative diffusion model”. *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2023, 12302–12311 1, 3.

[ZLH24] ZHANG, JIANFENG, LI, HENGXUAN, and HUO, ZHANQIANG. “Unsupervised Boosted Fusion Network for Single Low-light Image Enhancement”. *IEEE Access* (2024) 3.

[ZXW*21] ZHAO, ZUNJIN, XIONG, BANGSHU, WANG, LEI, et al. “RetinexDIP: A unified deep framework for low-light image enhancement”. *IEEE Transactions on Circuits and Systems for Video Technology* 32.3 (2021), 1076–1088 1, 2.