IniRetinex: Rethinking Retinex-type Low-Light Image Enhancer via Initialization Perspective

Guodong Fan[1,2], Zishu Yao[2], Guang-Yong Chen[3], Jian-Nan Su[3], Min Gan[2*]

[1]School of Compuger Science and Technotogy, Shandong Technology and Business University, Yantai, China
[2]College of Compuger Science and Technotogy, Qingdao University, Qingdao, China
[3]College of Computer and Data Science, Fuzhou University, Fuzhou, China
fgd96@outlook.com, yaozishu@qdu.edu.cn, gychen@fzu.edu.cn, sjn.fzu@gmail.com, aganmin@aliyun.com

Abstract

Retinex-based methods have become a general approach for solving low-light image enhancement (LLIE). However, traditional methods require post-processing of illumination (e.g., gamma correction), which lacks adaptability and disrupts the illumination structure. Retinex-based deep networks typically follow a 'decomposition-adjustment-exposure control' process, which is redundant and lacks robustness. One major issue is the inaccuracy in estimating and decomposing the initial illumination. Accurate initial illumination can prevent further post-processing instability. We propose IniRetinex, rethinking the Retinex-based LLIE method from the perspective of initialization. By using neural networks to provide reasonable initial illumination and solving for smooth illumination through optimization, higher performance LLIE is achieved. We construct a two-layer convolutional neural network to capture the low-frequency structure of the image, adaptively compensating for classical initial illumination and avoiding additional post-processing. The network requires no pre-training and can be implemented in an unsupervised manner with just a few iterations, making it highly efficient. Additionally, we propose a new illumination optimization strategy by introducing an additional proximal penalty term, improving illumination in areas with varying levels and enhancing image details. Extensive experiments on various low-light image datasets demonstrate that our method achieves state-of-the-art (SOTA) results on multiple benchmarks, offering higher stability and inference efficiency compared to current advanced methods.

Introduction

Image capture in real-world environments often faces numerous uncontrollable challenges that can degrade the quality and usefulness of the acquired visual data. Low-light conditions, such as nighttime scenes, backlit scenarios, or underexposed settings, serve as prime examples of such challenging scenarios (Li et al. 2021). These low-light environments can significantly impair the performance of advanced computer vision and image processing tasks, including image recognition, object detection and tracking, and image segmentation.

*Corresponding author
Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

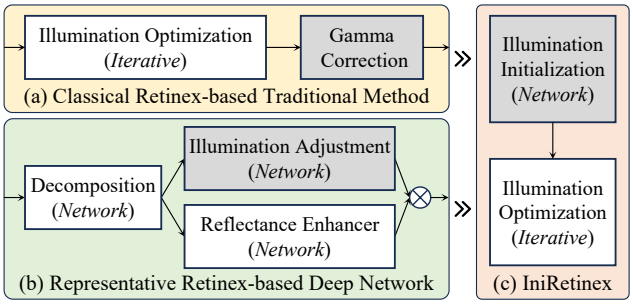

Figure 1: Classification of retinex-based methods and the general process of our method. (a) represents traditional methods that use numerical optimization to solve for illumination, often requiring post-processing like gamma correction. (b) shows the typical Retinex-based deep network structure, where initial illumination and reflectance are decomposed, optimized separately, and then combined to enhance the image. (c) illustrates our proposed InitRetinex, which focuses on obtaining accurate initial illumination through a neural network, followed by optimization. This approach retains the strengths of traditional methods while avoiding the drawbacks of post-processing.

Addressing the problems associated with low-light image capture is crucial for the reliable deployment of computer vision systems in real-world applications, where uncontrolled lighting conditions are frequently encountered. Developing robust techniques to enhance and recover information from low-quality, low-light images is an active area of research, with the goal of enabling advanced visual understanding even in challenging environmental conditions.

Related Works

Early low-light image enhancement methods used value mapping techniques like histogram equalization and gamma correction (Li et al. 2018a; Ma et al. 2023; Guo, Li, and Ling 2017). However, these methods focused on contrast enhancement without considering pixel context, leading to distortion and over-enhancement. Retinex theory (Land and McCann 1971), building on color constancy (Von Helmholtz 1867), became the mainstream approach for LLIE. Retinex-based methods are divided into traditional and deep network

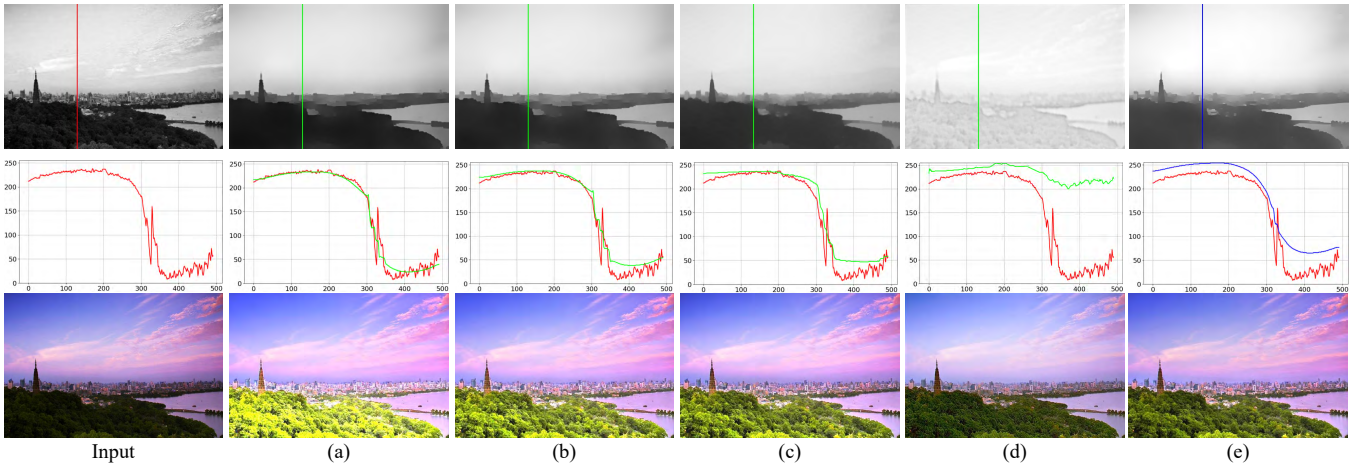

Figure 2: The impact of illumination obtained by different strategies on the enhancement results is analyzed. The top row shows the initial illumination and illumination generated by different strategies. The middle row presents line plots of pixel values along the color-marked regions in the illumination, used to compare the smoothness and variation in regional illumination. The bottom row displays the low-light input and the corresponding enhancement results, where (a) represents the illumination obtained directly using RTV, (b) applies gamma correction with a commonly used gamma value of 0.8 on (a), (c) shows the enhancement results from LIME (Guo, Li, and Ling 2017), (d) presents the enhancement results from URetinex (Wu et al. 2022), and (e) is our method.

approaches (as shown in Fig. 1). Traditional methods emphasize regularization, iterative illumination estimation, and enhancement through gamma correction. For example, Guo et al. (Guo, Li, and Ling 2016) used RTV (Xu et al. 2012) for illumination estimation, Hao et al. (Hao et al. 2020) used GTV-based filters for piecewise smooth illumination, and Cai et al. (Cai et al. 2017) proposed a joint intrinsic-extrinsic prior model for illumination and reflectance decomposition.

Deep network approaches (Wei et al. 2018; Zhang et al. 2021; Zhao et al. 2021; Jiang et al. 2022; Fan et al. 2022; Ju et al. 2023a,b) typically feature carefully designed architectures. These methods first decompose the low-light input into initial illumination and reflectance using a decomposition network, followed by separate networks for enhancing and denoising illumination and reflectance. Xu et al. (Xu et al. 2024) proposed a color shift-aware Retinex, which decomposes the image into reflectance, illumination, and color shift, correcting color biases for more realistic enhancement. Yang et al. (Yang et al. 2021) developed a sparse gradient minimization network to preserve major edge structures while removing low-amplitude information, effectively extracting illumination. Wu et al. (Wu et al. 2022) unfolded the optimization problem into a learnable network, using unfolding modules to adaptively learn illumination and reflectance. Cai et al. (Cai et al. 2023) first estimated illumination to brighten low-light images and then used a designed illumination-guided transformer to model non-local interactions in different lighting conditions. Given the limitations of learning from specific distributions, recent work (Zheng et al. 2023; Zhu et al. 2020; Fu et al. 2023a; Luo et al. 2023) has shifted towards unsupervised strategies to enhance generalization across different scenes. For instance, Zhang et al. (Zhang et al. 2020) developed a decomposition framework that simultaneously estimates illumination and reflectance. Sun et al. (Sun et al. 2024) introduced a digital imaging Retinex theory for computer vision, designing a brightness-contrast adjustment algorithm based on it. Liu et al. (Liu et al. 2021) designed a collaborative, no-reference learning strategy based on the Retinex principle to discover low-light priors from a compact search space. Ma et al. (Ma et al. 2022) employed a self-calibration module for illumination learning, but the complex constraint strategy struggled to learn smooth illumination, resulting in suboptimal performance. Although these Retinex-based methods achieve impressive enhancement quality, they often introduce complex learning mechanisms and constraints. When prior conditions are not met, it becomes challenging to learn accurate illumination, limiting the effectiveness of the enhancement.

Motivation

The main issue with traditional low-light image enhancement methods, such as gamma correction, is their inability to adaptively adjust illumination during post-processing. This often results in disruption of the illumination structure. Mainstream deep network approaches typically require a separate network to adjust illumination after decomposition, essentially a brightening-darkening process. This not only adds unnecessary computational cost but also makes the enhancement dependent on the training data distribution, resulting in poor robustness.

Analysis of different Retinex-based enhancements reveals the limitations of both traditional and deep learning methods (Fig. 2). Traditional methods usually require an initial illumination estimation, which can lead to overexposure issues due to fidelity constraints (Fig. 2(a)). Direct gamma correction can disrupt the overall illumination structure, degrading

visual quality (Fig. 2(b)). LIME (Guo, Li, and Ling 2017) also has similar issues (Fig. 2(c)). Deep network methods aim to adaptively find accurate illumination through training, but their performance is limited by the training data distribution, resulting in poor generalization and insufficient smoothness, causing detail loss (Fig. 2(d)).

The key challenge is to find a more reasonable initial illumination estimation, avoiding the shortcomings of both traditional and current deep learning approaches for low-light image enhancement.

Contribution

The current study focuses on the initial illumination estimation, which is the major issue of the limitations in existing LLIE approaches. We propose a method called IniRetinex to address the limitations of both traditional and deep learning approaches. The main contributions are:

- **Adaptive Initialization Strategy:** We propose a neural network-based approach to adaptively compensate for classical initial illumination estimation, achieving more accurate initial illumination without requiring additional training data.
- **Improved Illumination Optimization:** To mitigate the staircase effect caused by RTV-based illumination optimization, we introduce a proximal term that effectively weakens the illumination edge regions, resulting in smoother and more natural illumination estimation.
- **Enhanced Performance, Robustness, and Efficiency:** By addressing the key issues in initial illumination estimation and optimization, IniRetinex achieves higher performance, robustness, and computational efficiency compared to traditional and deep learning-based LLIE methods.

Method

Problem Definition

Retinex has become a foundational theory in the field of LLIE. It describes an image as the product of a reflectance layer, which is unaffected by changes in illumination, and an illumination layer:

$$\mathrm{L} = \mathrm{R} \circ \mathrm{I}, \tag{1}$$

where L represents the low-light image, R represents the image under normal illumination, and I represents the illumination degradation. The symbol $\circ$ denotes the Hadamard product.

In traditional methods, after obtaining the initial illumination $\hat{\mathrm{I}}$, the smoothness of the illumination is achieved by minimizing the following problem:

$$\min_{\mathrm{I}} \left\{ \mathrm{F}(\mathrm{I}, \hat{\mathrm{I}})) + \alpha\phi(\mathrm{I}) \right\}, \tag{2}$$

where F is the fidelity term used to preserve the image structure of I, ϕ is the regularization term ensuring the smoothness of I, and α represents the weight of this term. Edge-preserving smoothness methods are often employed in illumination optimization, with RTV (Xu et al. 2012) being a representative approach. Accordingly, the objective function is defined as:

$$\min_{\mathrm{I}} \|\mathrm{I} - \hat{\mathrm{I}}\|_2^2 + \lambda \sum_{*\in\{\mathrm{x},\mathrm{y}\}} \|\mathrm{W}_* \circ \partial_* \mathrm{I}\|_1, \tag{3}$$

where W is the weight matrix the initial illumination, The calculation method for W is referenced in ((Guo, Li, and Ling 2017)). $\hat{\mathrm{I}}$ is obtained by taking the maximum value across the RGB channels:

$$\hat{\mathrm{I}} = \mathcal{P}(\mathrm{L}) = \max_{c\in\{R,G,B\}} \mathrm{L}(c). \tag{4}$$

However, when RTV is used to solve for illumination, the fidelity term constrains the intensity of the illumination values, necessitating post-processing with gamma correction. This gamma correction, however, can disrupt the structure of the illumination obtained by RTV, resulting in unnatural detail rendering. Moreover, when RTV is applied to illumination estimation, it is prone to staircase effects due to inadequate protection of large gradients, leading to artifacts such as excessive edge smoothing and unnatural transitions.

To address these issues, we propose solutions from two perspectives: (1) adaptively modifying the initial value $\hat{\mathrm{I}}$; (2) introducing a proximal term tailored to the characteristics of the LLIE task. The optimization objective is therefore defined as follows:

$$\min_{\mathrm{I}} \|\mathrm{I} - \mathcal{T}_\theta(\mathrm{L})\|_2^2 + \sum_{*\in\{\mathrm{x},\mathrm{y}\}} (\eta\|\partial_*\mathrm{I} - \partial_*\hat{\mathrm{I}}\|_2^2 + \lambda\|\mathrm{W}_* \circ \partial_*\mathrm{I}\|_1), \tag{5}$$

where $\mathcal{T}_\theta$ represents the model designed in this paper to obtain the initial illumination from the input image, η and λ are regularization coefficients. To effectively overcome the staircase effect caused by insufficient protection of large gradients in the RTV method, we specifically designed penalty terms in the objective function: $\|\partial * \mathrm{I} - \partial_*\hat{\mathrm{I}}\|_2^2$ is used to protect large gradients, while $\|\mathrm{W}_* \circ \partial_*\mathrm{I}\|_1$ is used to enforce a sparse structure in the image gradients, thereby aiding in image smoothing. As η increases, the gradient of the output image approaches the gradient of the input image; as λ increases, the gradient of the output image becomes sparser,

Algorithm 1: IniRetinex

Input: A set of low-light images $\{\mathrm{L}_i\}_{i=1}^N$
for $i = 1, 2, \cdots, N$ do
  Step1: Adaptive initialization
  Preprocess L_i and initialize the network $\mathcal{G}$.
  for $j = 1, 2, \cdots, n$ do
    Compute the loss using Eq. (4) and Eq. (7).
    Update the parameters θ of $\mathcal{G}$.
  end for
  Step2: Illumination Optimization
  Obtain the initial illumination $\hat{\mathrm{I}}$ using Eq. (4) and Eq. (7).
  for $j = 1, 2, \cdots, k$ do
    Compute the illumination I_k using Eq. (14).
  end for
  Compute final R through Eq. (1).
end for
Output: Final enhanced image

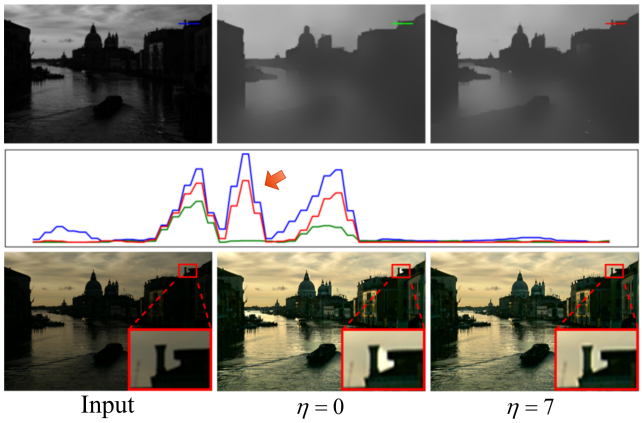

Figure 3: The top row displays the illumination, the middle row shows the gradient changes in the color-marked regions of the illumination, and the bottom row presents the low-light images alongside the enhanced images. The colored boxes highlight the zoomed-in image details.

Figure 4: The top row shows the low-light image and the outputs of $\mathcal{G}$ at 10, 30, 50, and 100 iterations. The middle row shows the illumination. The bottom row displays the enhanced images.

resulting in a smoother image. As shown in Fig. 3, unlike RTV, our method reduces deformation in regions with large gradients while smoothing the image, resulting in more natural edge transitions in the smoothed image. When η is 0, our method degenerates to the traditional RTV method.

Adaptive Initialization

Adaptive compensation of initial illumination (Forsyth 1990) is a challenging task. It requires adding different values to various regions of the initial illumination, which is inherently related to the illumination structure. Inspired by deep image priors (Ulyanov, Vedaldi, and Lempitsky 2018), CNNs inherently learn low-frequency structural information first during image reconstruction, demonstrating their higher sensitivity to these structures. We leveraged this characteristic to develop an Adaptive Initialization strategy:

$$\hat{\mathrm{I}} = \mathcal{T}_\theta(\mathrm{L}) = \mathcal{P}(\mathrm{L}) + \mathcal{G}_\theta(\mathrm{L}), \tag{6}$$

where $\mathcal{G}_\theta$ represents a convolutional neural network with learnable parameters θ, designed to capture low-frequency structures to compensate for the existing initial illumination. The structure of $\mathcal{G}$ as follow:

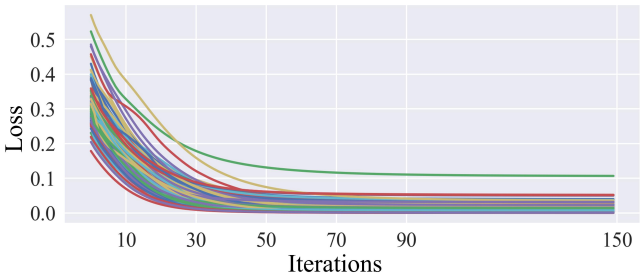

Figure 5: We observed the convergence of $\mathcal{G}$ by examining 100 separate examples.

$$\mathcal{G}(\mathrm{L}) = \mathrm{Sigmoid}(\mathrm{conv}_{8\to1}(\mathrm{GN}(\mathrm{conv}_{3\to8}(\mathrm{L})))). \tag{7}$$

We found that the constructed neural network to obtain adaptive compensation for the initial illumination is quite straightforward and can be achieved with a simple loss function:

$$\mathrm{Loss} = \mathrm{MSE}(\mathrm{L}, \hat{\mathrm{I}}). \tag{8}$$

As shown in Fig. 4 and 5, training $\mathcal{G}_\theta$ with loss function (8) will eventually drive $\mathcal{G}_\theta$ towards zero. However, this is not an issue, as the low-frequency structure can be obtained in very few iterations and used to effectively compensate for the initial illumination, producing excellent visual effects. This method is more robust than traditional gamma correction and requires minimal iterations to train $\mathcal{G}_\theta$, without relying on additional images or labels. Furthermore, we found that due to the simplified learning objective of $\mathcal{G}$, low-resolution images can be used during its training. Therefore, we train $\mathcal{G}$ using low-light images downsampled by a factor of 8 to accelerate the training process, while using a 2x downsampling during testing, which does not affect performance. The related experiments are provided in the ablation study.

Illumination Optimization

After obtaining the initial illumination $\mathcal{T}_\theta(\mathrm{L})$, we optimize the illumination by solving the established model (5). Eq. (5) is a composite optimization problem, which can be equivalently formulated as a constrained optimization problem:

$$\min_{\mathrm{I},\mathrm{z}_*} \|\mathrm{I} - \mathcal{T}_\theta(\mathrm{L})\|_2^2 + \sum_{*\in\{\mathrm{x},\mathrm{y}\}} (\eta\|\partial_*\mathrm{I} - \partial_*\hat{\mathrm{I}}\|_2^2 + \lambda\|\mathrm{z}_*\|_1), \tag{9}$$
$$s.t.,\ z_* = \mathrm{W}_* \circ \partial_*\mathrm{I}.$$

The corresponding augmented Lagrangian function can be written as:

$$\begin{aligned}\mathrm{L}(\mathrm{I}, z_*, u_*) = \|\mathrm{I} - \hat{\mathrm{I}}\|_2^2 + \sum_{*\in\{x,y\}} (\eta\|\partial_*\mathrm{I} - \partial_*\hat{\mathrm{I}}\|_2^2 + \lambda\|z_*\|_1) \\ + \frac{\rho}{2}\sum_{*\in\{x,y\}} (\|z_* - \partial_*\mathrm{I} + u_*\|_2^2 - \|u_*\|_2^2),\end{aligned} \tag{10}$$

where $u_{*\in\{x,y\}}$ are dual variables and $\rho > 0$ is the penalty parameter. Therefore, the illumination I can be iteratively

estimated using the Alternating Direction Method of Multipliers (ADMM) method:

$$\begin{aligned} \mathrm{I}^{k+1} = \underset{\mathrm{I}}{argmin} \, \|\mathrm{I} - \hat{\mathrm{I}}\|_2^2 + \eta \sum_{*\in\{x,y\}} \|\partial_* \mathrm{I} - \partial_* \hat{\mathrm{I}}\|_2^2 \\ + \frac{\rho}{2} \sum_{*\in\{x,y\}} \|z_*^k - \partial_* \mathrm{I} + u_*^k\|_2^2, \end{aligned} \tag{11}$$

$$z_*^{k+1} = prox_{\frac{\lambda}{\rho}\|\cdot\|_1}(\partial_* \mathrm{I}^{k+1} - u_*^k), \tag{12}$$

$$u_*^{k+1} = u_*^k + z_*^{k+1} - \partial_* \mathrm{I}^{k+1}. \tag{13}$$

Updating the illumination involves gradient operations, which typically require extensive matrix inversion calculations, imposing significant computational burdens that are sometimes unacceptable. This has prompted us to integrate the more efficient Fast Fourier Transform (FFT) to circumvent this challenge. FFT allows for computations to be conducted in the frequency domain using simpler element-wise multiplication, especially facilitating and expediting the process for systems with periodic boundary conditions. Specifically, with the aid of FFT and inverse FFT (iFFT) operations, the update of I can be analytically represented as:

$$\mathrm{I}^{k+1} = \mathcal{F}^{-1}\left(\frac{2\mathcal{F}(\hat{\mathrm{I}}) + \sum_{*\in\{x,y\}} \overline{\mathcal{F}(\partial_*)} \circ \mathcal{F}\left(2\eta\hat{\mathrm{I}} + \rho\left(\mathrm{z}_* + \mathrm{u}_*\right)\right)}{2\mathcal{F}(1) + (2\eta+\rho) \sum_{*\in\{x,y\}} \overline{\mathcal{F}(\partial_*)} \circ \mathcal{F}(\partial_*)}\right), \tag{14}$$

where $\mathcal{F}(\cdot)$ and $\mathcal{F}^{-1}(\cdot)$ are the FFT and IFFT operators, $\overline{\mathcal{F}(\cdot)}$ denotes the harmonic conjugate of $\mathcal{F}(\cdot)$, and $\mathcal{F}(1)$ is the FFT of the delta function. Additionally, Eq. (14) can be accelerated using GPU-based FFT computations, enabling efficient calculations.

In the updates of z_*^{k+1}- and u_*^{k+1}, we can utilize the over-relaxation technique to speed up the convergence process. Concretely, the quantity $\partial_* \mathrm{I}^{k+1}$ in Eq. (12) and Eq. (13) can be replaced with:

$$\beta\partial_* \mathrm{I}^{k+1} + (1-\beta) z_*^{k+1},$$

where β is an over-relaxed parameter. Mostly, $\beta \in [1.5, 1.8]$ is recommended. Based on experimental results, the parameters η, λ, ρ, k and β were configured to 7, 1, 15, 7 and 1.8, respectively.

Flexible Settings

In many real-world scenarios, noise is inevitably introduced during the LLIE process. Therefore, performing denoising steps is crucial. For low-light images with significant noise, a post-processing operation can be added as an enhanced method:

$$\mathrm{R} = \mathcal{N}_\Theta(\mathrm{L}./(\mathrm{I} + \epsilon)), \tag{15}$$

where $\mathcal{N}$ represents the denoising network for real-world scenarios, and Θ denotes the network parameters. ϵ is a small constant used to prevent division by zero. We employed an advanced network, specifically DANet (Yue et al. 2020). For DANet, we utilized its pre-trained model directly without retraining it on low-light data.

Experiments

In this section, we first introduce the details of the experiments and then thoroughly evaluate the proposed method. Additionally, we conduct an ablation study to examine and discuss the relevant parameters of the proposed method in detail.

Code — https://github.com/CCECfgd/IniRetinex

Experimental Details

Experimental Setup AdamW was used to train the $\mathcal{G}$ with a learning rate of 0.001. The code uses the PyTorch, and all experiments were conducted on a PC equipped with an Intel Core i7-13700K and an NVIDIA RTX 3090.

Compared Methods To evaluate the performance of the proposed method, we selected several representative approaches, including three traditional methods: LIME (Guo, Li, and Ling 2017) and RRM (Li et al. 2018b), as well as six of the latest deep network methods: Zero-DCE++ (Li, Guo, and Loy 2021), SCI (Ma et al. 2022), URetinex (Wu et al. 2022), PairLIE (Fu et al. 2023b), NeRCo (Yang et al. 2023) and ZRLLE-PQP (Wang et al. 2024).

Benchmarks and Metrics To evaluate the adaptability of the proposed method across various scenarios, we selected four classic no-reference benchmarks (LIME (Guo, Li, and Ling 2017), DICM (Lee, Lee, and Kim 2012), NPE(Wang et al. 2013) and Exdark (Loh and Chan 2019)), the reference benchmarks (LOL (Chen Wei 2018) and BAID (Lv et al. 2022)). For no-reference metrics, we employed Natural Image Quality Evaluator (NIQE) and Information entropy (IE). NIQE evaluates the naturalness and overall quality of the enhanced image by comparing its statistical features with a model of natural scenes. IE measures the richness of details and contrast in low-light enhanced images by assessing the distribution of pixel intensities. A low value of IE indicates insufficient brightness enhancement, suggesting that the image conveys less information. Conversely, a high value may imply excessive brightness, and therefore, it is necessary to combine visual effects for a comprehensive evaluation. For reference-based metrics, we selected the Structural Similarity Index (SSIM), Peak Signal-to-Noise Ratio (PSNR), Learned Perceptual Image Patch Similarity (LPIPS). Additionally, we calculated the number of parameters for different methods and their computational speed at 1080p resolution, which further contributes to a comprehensive evaluation of the methods.

Method Comparison and Analysis

Quantitative Evaluation As shown in Table 1, the proposed method demonstrates an advantage in the evaluation of NIQE and IE. In comparison with traditional methods, the proposed method achieves the highest values across all no-reference benchmarks for NIQE. However, it does not achieve the highest IE values because LIME excessively increases brightness in certain datasets. Although this results in the evaluation indicating more information, it reduces the

Methods	Publication	LIME		DICM		NPE		Exdark		LOLv1			BAID			Inference Times↓	Params↓
		NIQE↓	IE↑	NIQE↓	IE↑	NIQE↓	IE↑	NIQE↓	IE↑	SSIM↑	PSNR↑	LPIPS↓	SSIM↑	PSNR↑	LPIPS↓		
LIME	TIP·2017	4.075	7.608	3.465	7.353	3.598	7.543	3.852	7.516	0.731	17.879	0.217	0.786	16.991	0.125	2.307	/
RRM	TIP·2018	4.099	7.352	4.259	7.229	4.046	7.352	4.637	7.362	0.643	13.768	0.225	0.705	19.582	0.287	20.331	/
Zero-DCE++	TPAMI·2022	4.188	6.879	3.924	6.881	3.673	7.474	3.775	7.506	0.531	12.208	0.218	0.770	17.890	0.209	0.121	0.0106
SCI	CVPR·2022	4.165	7.202	3.662	7.076	3.912	7.307	3.975	7.438	0.528	13.806	0.253	0.738	16.656	0.159	0.025	0.0004
URetinex	CVPR·2022	4.390	7.504	3.850	7.406	3.883	7.358	4.189	7.455	0.822	20.141	0.079	0.854	20.511	0.215	0.314	0.3401
PairLIE	CVPR·2023	4.517	7.407	4.122	7.277	4.215	7.309	3.927	7.526	0.731	19.513	0.172	0.738	16.216	0.219	0.028	0.3418
NeRCo	ICCV·2023	4.007	7.268	4.052	7.284	4.030	7.355	4.092	7.323	0.791	23.081	0.102	0.707	18.899	0.301	2.083	7.8414
ZRLLE-PQP	CVPR·2024	5.070	7.262	4.768	7.288	4.248	7.292	4.625	7.216	0.773	18.791	0.134	0.714	17.267	0.303	/	1313.6
Ours		3.622	7.433	3.441	7.356	3.583	7.513	3.615	7.518	0.797	18.247	0.166	0.793	17.205	0.288	0.632 / 0.081	0.0003

Table 1: Quantitative evaluation on six benchmarks. Below, we present the number of parameters of the state-of-the-art LLIE methods, along with the inference time on images of 1080p resolution (the underline represents the inference time using the CPU). The best values are highlighted in red, and the second-best values are marked in blue.

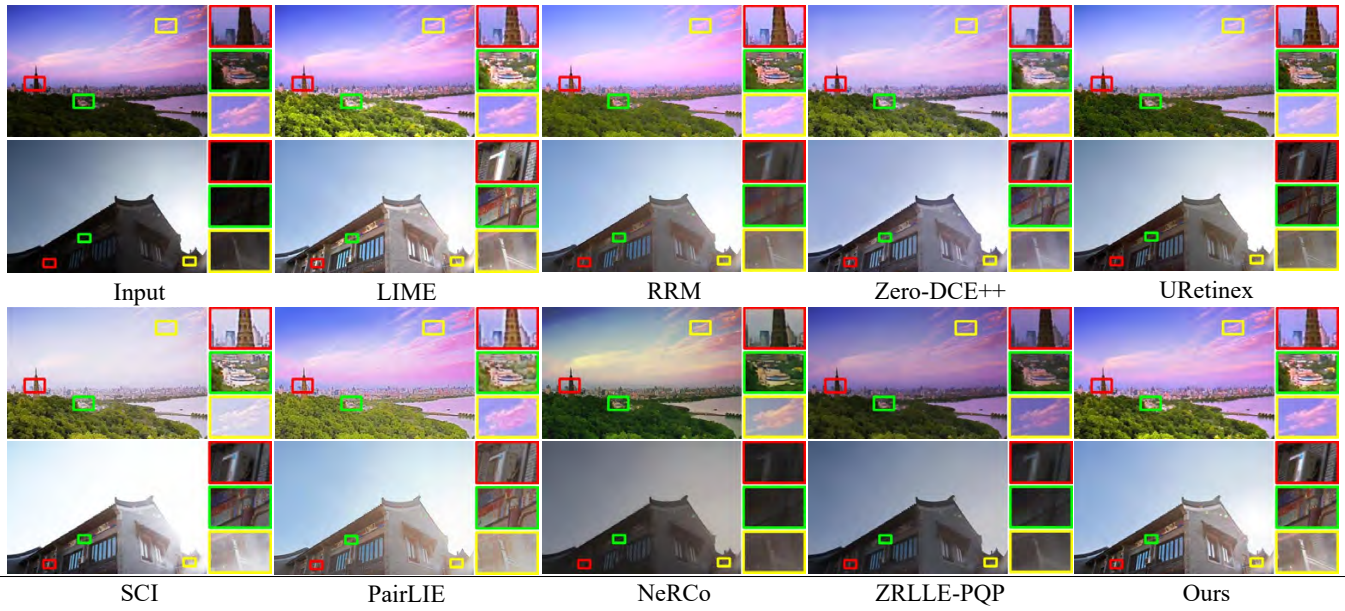

Figure 6: Comparison with current advanced methods in backlit scenes. The details within the color box are the extracted image details. The two examples are from NPE and BAID, respectively.

Figure 7: Comparison with current retinex-based methods in nighttime challenge scenes. The details within the color box are the extracted image details. The two examples are from LIME and Exdark, respectively.

visual quality of the images. In comparison with unsupervised methods, our use of neural networks, which are not affected by the distribution of the dataset, results in better performance across different types of no-reference data. This is also reflected in the comparison of NIQE and IE. For reference datasets, the methods trained on the LOL dataset are more sensitive to the distribution of LOL data, thus showing an advantage in reference evaluations. In the compari-

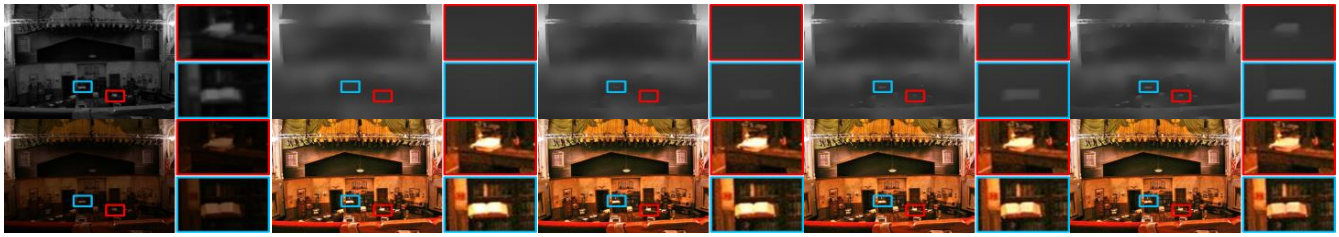

Figure 8: The value of η and its impact on enhancement quality. The first row shows the illumination, and the second row shows the enhanced results. The colored boxes highlight magnified local illumination and image details.

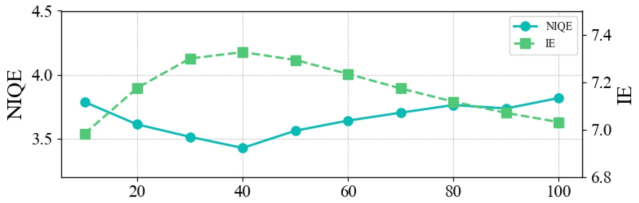

Figure 9: The impact of different iteration counts of $\mathcal{G}$ on performance.

Config	Size for training $\mathcal{G}$				Size for testing	
	1	↓2	↓4	↓8	↓2	↓4
NIQE↓	3.483	3.491	3.486	3.488	3.491	3.527
TIME↓	0.391	0.138	0.084	0.076	0.073	0.071

Table 2: Impact of different downsampling rates on performance and inference time during training and testing.

son of reference datasets, we compared traditional methods. In terms of inference time, our method shows significant improvement over traditional methods. However, it lags behind SCI and PairLIE because $\mathcal{G}$ requires individual training for each image, resulting in slightly longer inference times despite having fewer parameters.

Qualitative Evaluation In Fig. 6, the performance gap between methods in challenging uneven illumination scenarios is significant. LIME shows local overexposure due to gamma correction maintaining high exposure in bright areas but excessively increasing exposure in dark areas. RRM's brightness enhancement is insufficient, leading to unclear image details due to its objective function. Zero-DCE++, URetinex, and SCI alter the original image's tone to varying degrees. PairLIE brightens the image but its denoising process affects image details. NeRCo and ZRLLE-PQP fail to effectively enhance in such scenarios. Our method achieves the most natural visual effect and good detail restoration in this scenario.

Fig. 7 compares several mainstream Retinex-based methods in challenging nighttime scenarios. LIME significantly enhances brightness but amplifies compression artifacts in dark areas. URetinex shows varied performance across different examples, with color fading issues. SCI performs poorly in color restoration. PairLIE causes a smearing effect and amplifies compression artifacts. Our method not only significantly enhances image brightness but also achieves the most natural visual effect without noticeably amplifying compression artifacts in dark areas, demonstrating its advantages.

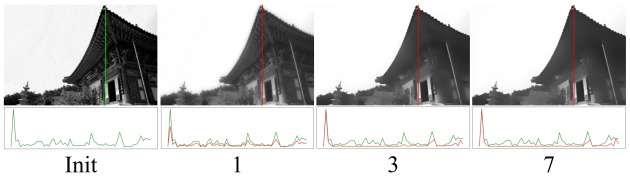

Figure 10: Comparison of smoothing performance with different iteration counts in illumination optimization. The first row shows the illumination, and the second row displays the gradient changes in the color-marked areas.

Ablation Study

About adaptive initialization We randomly selected 100 images from different datasets for an ablation study. We observed the performance of the method under different numbers of training iterations for $\mathcal{G}$, specifically ranging from 10 to 100 iterations. As shown in Fig. 9, the optimal value is achieved at 40 iterations. To accelerate inference speed, we applied bilinear interpolation downsampling to the input during the training of model $\mathcal{G}$. In Table 2, we first evaluated the impact of different downsampling rates during training, while keeping the input at full resolution during inference. The results show that using lower-resolution images during training has minimal impact on performance but significantly boosts speed. Consequently, we trained $\mathcal{G}$ with an 8x downsampled input and tested it with different downsampling rates. We found that while 2x downsampling during inference offered a good balance, 4x downsampling did not significantly improve speed but led to a substantial performance drop.

About illumination optimization We conducted parameter experiments on the added proximal term. As shown in Fig. 8, when η increases, issues like oversaturation and excessively bright areas are significantly alleviated because the gradients in detailed regions are effectively recaptured, indicating that the proximal term plays a role in preserving gra-

dients. We also experimented with the number of iterations for illumination optimization. As shown in Fig. 10, as the number of iterations increases, the extraneous details in the initial values are smoothed out, stabilizing by the seventh iteration.

Conclusion

In this paper, we focus on the initial illumination, achieving adaptive compensation for commonly used illumination through a simplified design that yields better visual performance than gamma correction. For illumination optimization, we identified the staircase effect of RTV when applied to LLIE tasks and found that adding a proximal term significantly mitigates this issue, preserving clearer illumination edges and thus protecting the enhanced image's edge regions. Compared to several traditional methods and popular unsupervised approaches, our method not only ranks among the top in quantitative evaluations but also delivers more natural visual results with higher computational efficiency.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under grant Nos. 6207308, 62173091, and 61772319; the Taishan Scholar Program of Shandong Province; and the Natural Science Foundation of Fujian Province under grant Nos. 2024109021 and 2023J01268.

References

Cai, B.; Xu, X.; Guo, K.; Jia, K.; Hu, B.; and Tao, D. 2017. A joint intrinsic-extrinsic prior model for retinex. In *Proceedings of the IEEE International Conference on Computer Vision*, 4000–4009.

Cai, Y.; Bian, H.; Lin, J.; Wang, H.; Timofte, R.; and Zhang, Y. 2023. Retinexformer: One-stage retinex-based transformer for low-light image enhancement. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 12504–12513.

Chen Wei, W. Y. J. L., Wenjing Wang. 2018. Deep Retinex Decomposition for Low-Light Enhancement. In *British Machine Vision Conference*.

Fan, G.-D.; Fan, B.; Gan, M.; Chen, G.-Y.; and Chen, C. L. P. 2022. Multiscale Low-Light Image Enhancement Network With Illumination Constraint. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(11): 7403–7417.

Forsyth, D. A. 1990. A novel algorithm for color constancy. *International Journal of Computer Vision*, 5(1): 5–35.

Fu, H.; Zheng, W.; Meng, X.; Wang, X.; Wang, C.; and Ma, H. 2023a. You do not need additional priors or regularizers in retinex-based low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 18125–18134.

Fu, Z.; Yang, Y.; Tu, X.; Huang, Y.; Ding, X.; and Ma, K.-K. 2023b. Learning a Simple Low-Light Image Enhancer From Paired Low-Light Instances. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 22252–22261.

Guo, X.; Li, Y.; and Ling, H. 2016. LIME: Low-light image enhancement via illumination map estimation. *IEEE Transactions on image processing*, 26(2): 982–993.

Guo, X.; Li, Y.; and Ling, H. 2017. LIME: Low-Light Image Enhancement via Illumination Map Estimation. *IEEE Transactions on Image Processing*, 26(2): 982–993.

Hao, S.; Han, X.; Guo, Y.; Xu, X.; and Wang, M. 2020. Low-light image enhancement with semi-decoupled decomposition. *IEEE Transactions on Multimedia*, 22(12): 3025–3038.

Jiang, K.; Wang, Z.; Wang, Z.; Chen, C.; Yi, P.; Lu, T.; and Lin, C.-W. 2022. Degrade is upgrade: Learning degradation for low-light image enhancement. In *Proceedings of the AAAI conference on artificial intelligence*, volume 36, 1078–1086.

Ju, Y.; Lam, K.-M.; Xiao, J.; Zhang, C.; Yang, C.; and Dong, J. 2023a. Efficient feature fusion for learning-based photometric stereo. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing*, 1–5. IEEE.

Ju, Y.; Zhang, C.; Huang, S.; Rao, Y.; and Lam, K.-M. 2023b. Learning deep photometric stereo network with reflectance priors. In *2023 IEEE International Conference on Multimedia and Expo*, 2027–2032. IEEE.

Land, E. H.; and McCann, J. J. 1971. Lightness and retinex theory. *Josa*, 61(1): 1–11.

Lee, C.; Lee, C.; and Kim, C.-S. 2012. Contrast enhancement based on layered difference representation. In *2012 19th IEEE International Conference on Image Processing*, 965–968.

Li, C.; Guo, C.; Han, L.; Jiang, J.; Cheng, M.-M.; Gu, J.; and Loy, C. C. 2021. Low-light image and video enhancement using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(12): 9396–9416.

Li, C.; Guo, C.; and Loy, C. C. 2021. Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(8): 4225–4238.

Li, M.; Liu, J.; Yang, W.; Sun, X.; and Guo, Z. 2018a. Structure-Revealing Low-Light Image Enhancement Via Robust Retinex Model. *IEEE Transactions on Image Processing*, 27(6): 2828–2841.

Li, M.; Liu, J.; Yang, W.; Sun, X.; and Guo, Z. 2018b. Structure-revealing low-light image enhancement via robust retinex model. *IEEE Transactions on Image Processing*, 27(6): 2828–2841.

Liu, R.; Ma, L.; Zhang, J.; Fan, X.; and Luo, Z. 2021. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10561–10570.

Loh, Y. P.; and Chan, C. S. 2019. Getting to know low-light images with the exclusively dark dataset. *Computer Vision and Image Understanding*, 178: 30–42.

Luo, Y.; You, B.; Yue, G.; and Ling, J. 2023. Pseudo-supervised low-light image enhancement with mutual learning. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(1): 85–96.

Lv, X.; Zhang, S.; Liu, Q.; Xie, H.; Zhong, B.; and Zhou, H. 2022. BacklitNet: A dataset and network for backlit image enhancement. *Computer Vision and Image Understanding*, 218: 103403.

Ma, L.; Liu, R.; Wang, Y.; Fan, X.; and Luo, Z. 2023. Low-Light Image Enhancement via Self-Reinforced Retinex Projection Model. *IEEE Transactions on Multimedia*, 25: 3573–3586.

Ma, L.; Ma, T.; Liu, R.; Fan, X.; and Luo, Z. 2022. Toward fast, flexible, and robust low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5637–5646.

Sun, S.; Ren, W.; Peng, J.; Song, F.; and Cao, X. 2024. DI-Retinex: Digital-Imaging Retinex Theory for Low-Light Image Enhancement. *arXiv preprint arXiv:2404.03327*.

Ulyanov, D.; Vedaldi, A.; and Lempitsky, V. 2018. Deep image prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 9446–9454.

Von Helmholtz, H. 1867. *Handbuch der physiologischen Optik*, volume 9. Voss.

Wang, S.; Zheng, J.; Hu, H.-M.; and Li, B. 2013. Naturalness Preserved Enhancement Algorithm for Non-Uniform Illumination Images. *IEEE Transactions on Image Processing*, 22(9): 3538–3548.

Wang, W.; Yang, H.; Fu, J.; and Liu, J. 2024. Zero-Reference Low-Light Enhancement via Physical Quadruple Priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

Wei, C.; Wang, W.; Yang, W.; and Liu, J. 2018. Deep retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*.

Wu, W.; Weng, J.; Zhang, P.; Wang, X.; Yang, W.; and Jiang, J. 2022. Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5901–5910.

Xu, H.; Zhang, H.; Yi, X.; and Ma, J. 2024. CRetinex: A Progressive Color-Shift Aware Retinex Model for Low-Light Image Enhancement. *International Journal of Computer Vision*, 1–23.

Xu, L.; Yan, Q.; Xia, Y.; and Jia, J. 2012. Structure extraction from texture via relative total variation. *ACM Transactions on Graphics*, 31(6): 1–10.

Yang, S.; Ding, M.; Wu, Y.; Li, Z.; and Zhang, J. 2023. Implicit neural representation for cooperative low-light image enhancement. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 12918–12927.

Yang, W.; Wang, W.; Huang, H.; Wang, S.; and Liu, J. 2021. Sparse gradient regularized deep retinex network for robust low-light image enhancement. *IEEE Transactions on Image Processing*, 30: 2072–2086.

Yue, Z.; Zhao, Q.; Zhang, L.; and Meng, D. 2020. Dual adversarial network: Toward real-world noise removal and noise generation. In *Proceedings of the European Conference on Computer Vision*, 41–58.

Zhang, Y.; Di, X.; Zhang, B.; and Wang, C. 2020. Self-supervised image enhancement network: Training with low light images only. *arXiv preprint arXiv:2002.11300*.

Zhang, Y.; Guo, X.; Ma, J.; Liu, W.; and Zhang, J. 2021. Beyond brightening low-light images. *International Journal of Computer Vision*, 129: 1013–1037.

Zhao, Z.; Xiong, B.; Wang, L.; Ou, Q.; Yu, L.; and Kuang, F. 2021. RetinexDIP: A unified deep framework for low-light image enhancement. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(3): 1076–1088.

Zheng, N.; Zhou, M.; Dong, Y.; Rui, X.; Huang, J.; Li, C.; and Zhao, F. 2023. Empowering low-light image enhancer through customized learnable priors. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 12559–12569.

Zhu, A.; Zhang, L.; Shen, Y.; Ma, Y.; Zhao, S.; and Zhou, Y. 2020. Zero-shot restoration of underexposed images via robust retinex decomposition. In *2020 IEEE International Conference on Multimedia and Expo*, 1–6. IEEE.