

Article

SCAUNet: Step-Size-Consistent ADMM Unfolding Network for Low-Light Image Enhancement

Xiaofang Li ¹ , Hongbiao Tian ¹ and Cui Fu ^{2,*}¹ College of Science, Northwest A&F University, Xianyang 712100, China; xfli@nwafu.edu.cn (X.L.); 2023014517@nwafu.edu.cn (H.T.)² School of Mathematics and Information Science, Xianyang Normal University, Xianyang 712000, China

* Correspondence: cuifu@stu.xidian.edu.cn

Abstract

Low-light image enhancement aims to restore visually pleasing normal-light images from degraded low-light observations. Most existing methods handle luminance variation from the enhancement perspective. As a result, the degradation process from a normal-light image to a low-light observation is usually not explicitly characterized. In addition, degradation-oriented optimization is often computationally expensive due to repeated iterative updates. To address these issues, based on the alternating direction method of multipliers (ADMM), a degradation-oriented step-size-consistent unfolding network SCAUNet is proposed. Specifically, a low-light image is modeled as the element-wise product of a normal-light image and a luminance degradation operator, together with additive noise. Based on this formulation, low-light enhancement is converted into the joint estimation of the target image and the degradation operator. Then, a state-based one-step ADMM solver is developed, and a step-size consistency constraint is introduced to improve the reliability of one-step unfolding. Extensive experiments on LOL-v1 and LOL-v2 demonstrate the effectiveness of the proposed SCAUNet. Compared with existing state-of-the-art methods, SCAUNet yields better enhancement quality, especially in preserving image structures, correcting illumination, and suppressing artifacts. Strong generalization ability is also verified on four no-reference low-light datasets, and promising results are obtained on single-image exposure correction.

Keywords: low-light image enhancement; unfolding network; step-size consistency; degradation modeling; exposure correction

MSC: 68U10; 68T07



Academic Editor: Andreas Weinmann

Received: 25 April 2026

Revised: 29 May 2026

Accepted: 8 June 2026

Published: 9 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Low-light images are commonly captured in scenes with insufficient illumination or limited exposure time, where the imaging process often suffers from severe photon deficiency and unstable acquisition conditions. As a result, the captured images usually exhibit reduced visibility, low contrast, and amplified noise, which significantly degrade perceptual quality and structural fidelity [1]. Such degradations not only impair human visual experience, but also undermine the reliability of subsequent high-level vision tasks, such as object detection, tracking, surveillance, autonomous driving, and computational photography [2,3]. Therefore, low-light image enhancement (LLIE) aims to restore a visually clear and semantically faithful normal-light image from a degraded low-light observation, and has become a fundamental problem in low-level vision.

At the core of low-light image enhancement lies the problem of how to model and correct the luminance variation from low-light observations to normal-light images. Existing deep learning based LLIE methods differ in how they interpret and handle this luminance transition process. From this perspective, deep learning-based LLIE methods roughly fall into four categories: direct mapping methods, structure–texture decomposition (STD)-based methods, Retinex-based methods and transfer function-based methods.

The first line of work treats enhancement as direct mapping from low-light images to normal-light images, where deep neural networks are trained in an end-to-end manner to implicitly learn the brightness correction process. Representative implementations in this category span a wide range of architectures, including transformer-based models [4], flow-based models [5,6], mamba-based models [7], and more recently generative paradigms [8,9]. While these methods are flexible and often achieve strong empirical performance, the luminance variation is largely absorbed into a black box transformation. This makes it difficult to explicitly disentangle illumination attenuation from noise corruption and to impose physically meaningful constraints on the enhancement process.

A second line of work models luminance variation through structure–texture decomposition. These methods first decompose a low-light image into a structure layer and a texture layer. Such decomposition is usually performed in the spatial domain [10,11], feature domain [12], or wavelet domain [13]. The luminance correction process is then mainly performed on the structure layer, while the texture layer is handled by denoising or detail restoration. Compared with direct enhancement methods, this formulation introduces stronger structural priors and improves interpretability. However, the luminance transition is still modeled on an intermediate decomposed representation. It is not derived from an explicit low-light degradation process. In addition, the enhancement pipeline usually consists of decomposition, separate processing, and reconstruction, which increases the overall complexity.

Another important line of work is based on Retinex theory. These methods decompose a low-light image into reflectance and illumination, and treat luminance variation mainly as illumination change. One group uses the estimated illumination to derive an initial reflectance image, and then directly restores it to obtain the final enhanced result. Representative methods include DeepUPE [14], RUAS [15], Retinexformer [16], and Retinexmamba [17]. The other group enhances the illumination component and restores the reflectance component separately, and then combines them to reconstruct the enhanced image. Early methods, such as RetinexNet [18], KinD [19], and KinD++ [20], mainly rely on dedicated Retinex decomposition, illumination adjustment and reflectance restoration modules. More recent methods further improve this framework by incorporating optimization-inspired unfolding [21–24] into the Retinex decomposition module or introducing diffusion-based modeling [25] for stronger restoration capability. This line of work is more interpretable than direct enhancement methods and STD-based methods. However, most existing methods still handle luminance transition through illumination enhancement or reflectance restoration. They usually rely on heuristic priors or data-driven mappings, rather than explicitly modeling the low-light luminance degradation process.

To simplify the decomposition pipeline, several studies have modeled luminance variation through a transfer function [26,27]. In these methods, brightness correction is usually formulated as the estimation of a transfer function that maps a low-light image to a coarse normal-light image, followed by an additional refinement module. Such formulations are closer to explicit luminance correction than direct black box enhancement. However, the transfer function is mainly introduced from the enhancement perspective. It is used to transform a low-light image into a brighter one, rather than to characterize the degradation process from a normal-light image to a low-light observation. Conse-

quently, the luminance variation is not explicitly incorporated into a low-light image formation model. This limits the construction of a unified framework in which the target normal-light image and the luminance degradation process can be jointly estimated.

A related attempt was made in CERL [28], where a low-light image is modeled as the result of sequential noise corruption and luminance attenuation applied to a normal-light image. This formulation is closer to the degradation perspective. However, the luminance degradation process is not represented by an explicit operator to be estimated. Instead, the restoration pipeline still follows two separate stages, i.e., image brightening and noise removal. As a result, the target normal-light image is not recovered through the joint estimation of the image and the luminance degradation operator. A degradation-oriented formulation is therefore still required.

Although degradation modeling makes low-light enhancement more interpretable, the resulting optimization problem is still difficult to solve efficiently. The target normal-light image and the luminance degradation operator are coupled with each other, and their estimation usually relies on iterative optimization with repeated alternating updates. The alternating direction method of multipliers (ADMM) offers a suitable framework for handling such a problem, since it enables variable splitting and subproblem-based optimization. However, a standard ADMM solver requires multiple iterations, and its unfolded multi-stage counterpart still suffers from high computational and memory cost. A one-step unfolding scheme is therefore appealing for efficient enhancement. However, once the iterative process is compressed into a single update, the gradual correction property of the original solver can be weakened. The learned update may become unstable or overly strong, and the consistency between one large step and multiple smaller steps is no longer guaranteed. This makes it necessary to introduce an additional mechanism to regularize the behavior of one-step unfolding.

Motivated by the above limitations, low-light enhancement is formulated from the degradation perspective. Specifically, a low-light image is modeled as the element-wise product of a normal-light image and a luminance degradation operator, together with additive noise. Under this formulation, luminance variation is explicitly embedded into the low-light image formation process, rather than treated as heuristic enhancement mapping. Then, a state-based ADMM-based framework is introduced to jointly estimate the target normal-light image and the luminance degradation operator. To reduce the cost of repeated iterative updates, the solver is further unfolded into a one-step network. To preserve the progressive behavior of iterative optimization, a step-size consistency constraint is imposed, such that one large update is constrained to be consistent with two smaller updates. Although the proposed framework is mainly developed for low-light image enhancement, the underlying formulation is not limited to under-exposed inputs. Since the luminance variation is modeled through an explicit operator, the same framework can also be extended to more general single-image exposure correction.

In summary, the contributions of this paper are four-fold:

- A new degradation modeling framework is presented for low-light image enhancement, in which the low-light image formation process is explicitly characterized by a luminance degradation operator.
- A state-based ADMM-inspired joint estimation framework is constructed and further unfolded into a one-step network. This design preserves interpretability while avoiding the high cost of multi-stage iterative inference.
- A step-size consistency constraint is introduced to regularize one-step unfolding. This strategy enforces the consistency between one large update and two smaller updates, leading to a more stable and efficient one-step unfolding process.

- Extensive experiments verify the effectiveness of the proposed method on low-light image enhancement and its generalization to single-image exposure correction.

The rest of this paper is organized as follows. Some closely related works are reviewed in Section 2. Then the proposed SCAUNet is described in Section 3. Experimental results and analysis are provided in Sections 4 and 5. Conclusions are finally drawn in Section 6.

2. Related Work

Existing deep learning-based LLIE methods can be roughly grouped into direct mapping methods, structure–texture decomposition based methods, Retinex-based methods, and transfer-function-based methods.

Direct mapping methods learn black box mapping from low-light images to normal-light images in an end-to-end manner. Representative methods mainly differ in the network architecture. For example, SNR-Net [4] introduces signal-to-noise priors into the network, LLFlow [5] and UPTFlow [6] adopt flow-based formulations, MambaLLIE [7] is built on the Mamba architecture, and LightenDiffusion [8] and DiffLLE [9] further explore diffusion-based generation. Despite their effectiveness, the underlying luminance attenuation and noise corruption are not explicitly modeled.

STD-based methods first decompose a low-light image into structure and texture components, and then enhance them separately. Representative methods differ mainly in the domain where the decomposition is performed. For example, some methods conduct structure–texture decomposition in the spatial domain [10,11], while STANet [12] performs this process in the feature domain, and DiffLL [13] further introduces decomposition in the wavelet domain. Under this formulation, luminance correction is mainly applied to the structure component, while the texture component is restored separately. Although such methods improve interpretability, the luminance variation is still handled on an intermediate decomposed representation, rather than through an explicit low-light degradation process.

Retinex-based methods model low-light enhancement through illumination–reflectance decomposition, and attribute luminance variation mainly to illumination change. One group of methods estimates the illumination component and derives the enhanced image from the corresponding reflectance. For example, DeepUPE [14] predicts the illumination map and restores the image by illumination-guided inversion, while RUAS [15] formulates enhancement within a Retinex-inspired unrolling framework. Retinexformer [16] and Retinexmamba [17] further estimate the illumination component, derive an initial reflectance image, and then restore it by different backbone architectures. Another group of methods jointly enhance illumination and restore reflectance. This line of work originates from RetinexNet [18], and is further extended by KinD [19] and KinD++ [20] with dedicated reflectance restoration modules. More recently, URetinex-Net [21], URetinex-Net++ [22], RAUNA [23], and CUE [24] incorporate unfolding-based optimization into Retinex modeling, while Diff-Retinex++ [25] combines diffusion modeling with the Retinex framework. Although these methods are more interpretable, the luminance transition is still mainly handled through illumination refinement or reflectance restoration, rather than through an explicit degradation process.

Transfer function-based methods model luminance variation through an explicit transfer function. TBEFN [26] models a normal-light image as the element-wise product of a clean low-light image and a transfer function. SATNet [27] further estimates the transfer function with the structure pairs. This formulation is closer to explicit luminance correction, since the brightness change is represented by an operator. However, f is introduced from the enhancement perspective. It describes how a low-light image is transformed into a normal-light image, rather than how a normal-light image degrades into a low-light observation.

A further step is taken in CERL [28], where low-light imaging is modeled as a two-stage degradation process as follows:

$$\mathbf{O}_n = \arg \min_{\mathbf{O}_n} \frac{1}{2} \|\mathbf{O}_l - L(N(\mathbf{O}_n))\|_2^2 + \lambda \Phi(\mathbf{O}_n), \quad (1)$$

where $(\mathbf{O}_l, \mathbf{O}_n)$ denote the low-light and the corresponding normal-light images, respectively. $L(\cdot)$ is the illumination attenuation function, $N(\cdot)$ represents the noise corruption function, $\Phi(\cdot)$ denotes the regularization term of $\mathbf{O}_n$, and λ is a tuning parameter. This formulation moves closer to degradation modeling. However, the attenuation function $L(\cdot)$ remains implicit. Moreover, the restoration pipeline still follows two separate stages, i.e., image brightening and noise removal.

3. The Proposed Method

Existing low-light enhancement methods usually handle luminance variation from the enhancement perspective. As a result, the degradation process from a normal-light image to a low-light observation is often not explicitly characterized. Moreover, although degradation-oriented formulations are more interpretable, their optimization usually relies on repeated iterative updates, which lead to high computational cost. Motivated by these limitations, the proposed method is developed from the degradation perspective. A low-light image is modeled as the element-wise product of a normal-light image and a luminance degradation operator, together with additive noise. Based on this formulation, low-light enhancement is converted into the joint estimation of the target normal-light image and the degradation operator. To solve this problem efficiently, a state-based one-step ADMM solver is constructed and further unfolded into a network. To improve the reliability of the proposed one-step unfolding network, a step-size consistency constraint is introduced. The overall framework is shown in Figure 1. The degradation model, the state-based one-step ADMM solver, and the step-size consistency constraint are presented in the following subsections.

3.1. Degradation Model

In this paper, the low-light image formation process is modeled as

$$\mathbf{O}_l = \mathbf{O}_n \cdot \mathbf{G} + \mathbf{N}, \quad (2)$$

where $\mathbf{G}$ is the luminance degradation operator and $\mathbf{N}$ is additive noise. Similar to Retinex-based and transfer function-based formulations, luminance variation is described by an explicit multiplicative component. The difference is that the proposed model introduces this component from the degradation perspective, rather than from the enhancement perspective. Therefore, low-light enhancement is formulated as the joint estimation of the target normal-light image $\mathbf{O}_n$ and the degradation operator $\mathbf{G}$ from the observation $\mathbf{O}_l$.

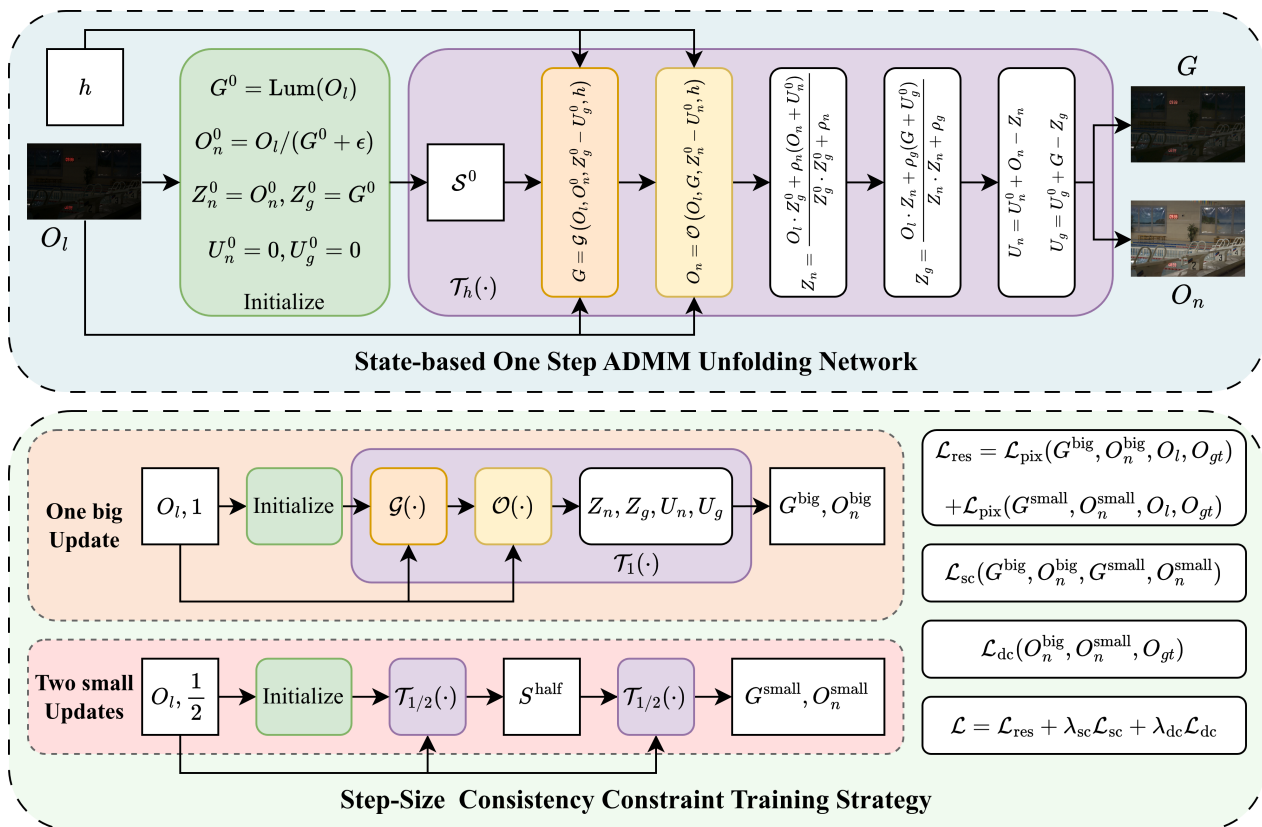


Figure 1. Illustration of the proposed SCAUNet. The upper part presents the state-based one-step ADMM unfolding network with physically meaningful initialization and one-step state transition. The lower part shows the step-size consistency constraint training strategy, where one large step update is regularized by two successive half step updates.

3.2. State-Based One-Step ADMM Solver

Based on Equation (2), the restoration of O_n and G is formulated as

$$\min_{O_n, G} \frac{1}{2} \|O_l - O_n \cdot G\|_2^2 + \lambda_n \Phi_n(O_n) + \lambda_g \Phi_g(G), \quad (3)$$

where $\Phi_n(\cdot)$ and $\Phi_g(\cdot)$ denote the regularization terms of O_n and G , respectively. λ_g and λ_n represent their corresponding regularization weights. In this paper, these regularization terms are not implemented in explicit analytical form, but are instead modeled implicitly by learnable solvers. Therefore, λ_g and λ_n are introduced here to motivate the degradation-aware optimization formulation, rather than being explicitly tuned in the actual implementation. Then, ADMM is introduced as a variable splitting strategy to decouple the data fidelity term and the implicit regularization terms. However, directly unfolding many ADMM iterations would lead to higher model complexity. Therefore, a one-step unfolding design is adopted. Let Z_n and Z_g denote the auxiliary variables associated with O_n and G , respectively, and let U_n and U_g denote the corresponding Lagrangian multipliers. After variable splitting, O_n , G , Z_n , Z_g , U_n , and U_g are jointly updated in one ADMM iteration. They are grouped into a unified state representation as

$$\mathcal{S} = \{O_n, G, Z_n, Z_g, U_n, U_g\}. \quad (4)$$

Therefore, one unfolding step corresponds to one ADMM update of the state $\mathcal{S}$.

Since the proposed framework unfolds only one ADMM step, the quality of the initial state is particularly important. To reduce the gap between the initialized variables and the

underlying optimum, a physically meaningful initial state is constructed from the observed low-light image as follows:

$$\mathbf{G}^0 = \text{Lum}(\mathbf{O}_l), \quad \mathbf{O}_n^0 = \mathbf{O}_l / (\mathbf{G}^0 + \epsilon), \quad \mathbf{Z}_n^0 = \mathbf{O}_n^0, \quad \mathbf{Z}_g^0 = \mathbf{G}^0, \quad \mathbf{U}_n^0 = \mathbf{0}, \quad \mathbf{U}_g^0 = \mathbf{0}, \quad (5)$$

where ϵ is a small constant. $\text{Lum}(\cdot)$ denotes a luminance extraction operator used to initialize $\mathbf{G}^0$, which is defined as

$$\text{Lum}(\mathbf{O}_l) = \text{Repeat}\left(\text{AvgPool}\left(0.299\mathbf{O}_l^{(R)} + 0.587\mathbf{O}_l^{(G)} + 0.114\mathbf{O}_l^{(B)}\right)\right), \quad (6)$$

where $\text{AvgPool}(\cdot)$ denotes local average smoothing and $\text{Repeat}(\cdot)$ replicates the luminance map along the channel dimension. In this way, $\mathbf{G}^0$ provides an illumination-related estimate from the luminance of the observation, and $\mathbf{O}_n^0$ gives a coarse estimate of the normal-light image. The remaining variables are then initialized consistently with this estimate, yielding an ADMM state that is already physically plausible before the one-step update.

Starting from $\mathcal{S}^0 = \{\mathbf{O}_n^0, \mathbf{G}^0, \mathbf{Z}_n^0, \mathbf{Z}_g^0, \mathbf{U}_n^0, \mathbf{U}_g^0\}$, the variables $\mathbf{G}$, $\mathbf{O}_n$, $\mathbf{Z}_n$, $\mathbf{Z}_g$, $\mathbf{U}_n$, and $\mathbf{U}_g$ are updated by the following one-step state transition. The updates of $\mathbf{O}_n$ and $\mathbf{G}$ are written as the following two subproblems:

$$\mathbf{G} = \arg \min_{\mathbf{G}} \lambda_g \Phi_g(\mathbf{G}) + \frac{\rho_g}{2} \|\mathbf{G} - \mathbf{Z}_g^0 + \mathbf{U}_g^0\|_2^2, \quad (7)$$

$$\mathbf{O}_n = \arg \min_{\mathbf{O}_n} \lambda_n \Phi_n(\mathbf{O}_n) + \frac{\rho_n}{2} \|\mathbf{O}_n - \mathbf{Z}_n^0 + \mathbf{U}_n^0\|_2^2. \quad (8)$$

Then, two learnable solvers are introduced to directly estimate their solutions as follows:

$$\mathbf{G} = \mathcal{G}(\mathbf{O}_l, \mathbf{O}_n^0, \mathbf{Z}_g^0 - \mathbf{U}_g^0, h), \quad (9)$$

$$\mathbf{O}_n = \mathcal{O}(\mathbf{O}_l, \mathbf{G}, \mathbf{Z}_n^0 - \mathbf{U}_n^0, h), \quad (10)$$

where $\mathcal{O}(\cdot)$ and $\mathcal{G}(\cdot)$ denote two learnable solvers, and h is the step size. Moreover, ρ_g and ρ_n are set as learnable parameters and optimized jointly with the network. The network $\mathcal{G}(\cdot)$ adopts a lightweight three-layer structure, consisting of two 1×1 convolutions and a 5×5 depth-wise-separable convolution to efficiently model spatial luminance variations. The network $\mathcal{O}(\cdot)$ follows the mamba-based architecture [17] to capture long-range dependencies.

After $\mathbf{O}_n$ and $\mathbf{G}$ are obtained, the remaining variables are updated analytically as

$$\mathbf{Z}_n = (\mathbf{O}_l \cdot \mathbf{Z}_g^0 + \rho_n(\mathbf{O}_n + \mathbf{U}_n^0)) / (\mathbf{Z}_g^0 \cdot \mathbf{Z}_g^0 + \rho_n), \quad (11)$$

$$\mathbf{Z}_g = (\mathbf{O}_l \cdot \mathbf{Z}_n + \rho_g(\mathbf{G} + \mathbf{U}_g^0)) / (\mathbf{Z}_n \cdot \mathbf{Z}_n + \rho_g), \quad (12)$$

$$\mathbf{U}_n = \mathbf{U}_n^0 + \mathbf{O}_n - \mathbf{Z}_n, \quad (13)$$

$$\mathbf{U}_g = \mathbf{U}_g^0 + \mathbf{G} - \mathbf{Z}_g. \quad (14)$$

The proposed state-based one-step ADMM solver is therefore written as

$$\mathcal{S} = \mathcal{T}_h(\mathcal{S}^0; \mathbf{O}_l), \quad (15)$$

where $\mathcal{T}_h(\cdot)$ denotes one ADMM update with step size h . The whole procedure is summarized in Algorithm 1.

Algorithm 1 The proposed state-based one-step ADMM algorithm**Input:** The observed low-light image O_l .

1: Compute the initial state by

$$G^0 = \text{Lum}(O_l), \quad O_n^0 = O_l / (G^0 + \epsilon), \quad Z_n^0 = O_n^0, \quad Z_g^0 = G^0, \quad U_n^0 = 0, \quad U_g^0 = 0.$$

2: Estimate the luminance degradation operator by

$$G = \mathcal{G}(O_l, O_n^0, Z_g^0 - U_g^0, h).$$

3: Estimate the normal-light image by

$$O_n = \mathcal{O}(O_l, G, Z_n^0 - U_n^0, h).$$

4: Update the auxiliary variables Z_n and Z_g by

$$Z_n = (O_l \cdot Z_g^0 + \rho_n(O_n + U_n^0)) / (Z_g^0 \cdot Z_g^0 + \rho_n),$$

$$Z_g = (O_l \cdot Z_n + \rho_g(G + U_g^0)) / (Z_n \cdot Z_n + \rho_g).$$

5: Update the Lagrangian multipliers by

$$U_n = U_n^0 + O_n - Z_n, \quad U_g = U_g^0 + G - Z_g.$$

Output: G, O_n .**3.3. Step-Size Consistency Constraint**

A one-step update is efficient, but it may become overly aggressive. As a result, the consistency between one large update and multiple smaller updates may be weakened. To alleviate this issue, a step-size consistency constraint is introduced. Let $\mathcal{T}_h(\cdot; O_l)$ denote the proposed state-based one-step ADMM solver with step size h , where O_l is the fixed low-light observation. Given the initial state $\mathcal{S}^0$, one large step update is written as

$$\mathcal{S}^{\text{big}} = \mathcal{T}_1(\mathcal{S}^0; O_l), \quad (16)$$

while two successive small-step updates are written as

$$\mathcal{S}^{\text{half}} = \mathcal{T}_{1/2}(\mathcal{S}^0; O_l), \quad (17)$$

$$\mathcal{S}^{\text{small}} = \mathcal{T}_{1/2}(\mathcal{S}^{\text{half}}; O_l). \quad (18)$$

The proposed constraint requires these two update paths to produce consistent final states:

$$\mathcal{T}_1(\mathcal{S}^0; O_l) \approx \mathcal{T}_{1/2}(\mathcal{T}_{1/2}(\mathcal{S}^0; O_l); O_l). \quad (19)$$

In the proposed implementation, the step size is introduced into the learnable solvers for G and O_n . Therefore, different step sizes lead to different estimates of the main variables, which further affect the subsequent analytical updates of Z_n , Z_g , U_n , and U_g . Based on Equation (19), the step-size consistency loss is defined as

$$\mathcal{L}_{\text{sc}} = \|O_n^{\text{big}} - O_n^{\text{small}}\|_1 + \|G^{\text{big}} - G^{\text{small}}\|_1, \quad (20)$$

where O_n^{big} and O_n^{small} are the restored normal-light images read out from $\mathcal{S}^{\text{big}}$ and $\mathcal{S}^{\text{small}}$, respectively. G^{big} and G^{small} are the luminance degradation operators read out from $\mathcal{S}^{\text{big}}$ and $\mathcal{S}^{\text{small}}$, respectively.

For training, both the one-step branch and the two-half-step branch are supervised. In the latter branch, the first half-step is fixed, and only the second half-step is learnable. Given an estimated normal-light image O_n , reference normal-light image O_{gt} , and degradation operator G , the pixel loss is defined as

$$\mathcal{L}_{\text{pix}} = \|O_n - O_{gt}\|_1 + 1 - \text{SSIM}(O_n, O_{gt}) + \|G \cdot O_{gt} - O_l\|_1 + 1 - \text{SSIM}(G \cdot O_{gt}, O_l), \quad (21)$$

where $\text{SSIM}(\cdot)$ is the structural similarity (SSIM) metric [29]. Accordingly, the restoration loss is written as

$$\mathcal{L}_{\text{res}} = \mathcal{L}_{\text{pix}}(G^{\text{big}}, O_n^{\text{big}}, O_l, O_{gt}) + \mathcal{L}_{\text{pix}}(G^{\text{small}}, O_n^{\text{small}}, O_l, O_{gt}). \quad (22)$$

To further improve color vividness and local contrast, a dark channel loss is introduced as an auxiliary supervision term. It is defined as

$$\mathcal{L}_{\text{dc}} = \left\| \mathcal{D}(O_n^{\text{big}}) - \mathcal{D}(O_{gt}) \right\|_1 + \left\| \mathcal{D}(O_n^{\text{small}}) - \mathcal{D}(O_{gt}) \right\|_1, \quad (23)$$

where $\mathcal{D}(\cdot)$ is utilized to extract the dark channel of a restored image, which is defined as

$$\mathcal{D}(O_n)(x) = \min_{y \in \Omega(x)} \left(\min_{c \in \{R, G, B\}} O_n^{(c)}(y) \right),$$

where $\Omega(x)$ denotes a local neighborhood centered at pixel x .

Together with the step-size consistency loss, the overall training objective is defined as

$$\mathcal{L} = \mathcal{L}_{\text{res}} + \lambda_{\text{sc}} \mathcal{L}_{\text{sc}} + \lambda_{\text{dc}} \mathcal{L}_{\text{dc}}. \quad (24)$$

4. Experiments

4.1. Experimental Settings

4.1.1. Datasets

Experiments are conducted on both paired and unpaired low-light image datasets. For supervised evaluation, the proposed method is evaluated on LOL-v1 [18] and LOL-v2 [30]. LOL-v1 [18] contains 500 paired low-/normal-light images, among which 485 pairs are used for training and the remaining 15 pairs are used for testing. LOL-v2 [30] contains 689 training pairs and 100 testing pairs of real low-/normal-light images. Due to the limited number of test images in LOL-v1 and LOL-v2, the GT mean method [19] is utilized during testing. To further examine the generalization ability of the proposed framework, the model trained on LOL-v1 is directly applied to no-reference low-light image datasets without additional fine-tuning. These datasets include LIME [31], DICM [32], MEF [33], and NPE [34], with 10, 65, 17, and 64 images, respectively. In addition, the proposed SCAUNet can be extended to more general single-image exposure correction. The corresponding experiments are conducted on SICE [35]. SICE contains 589 sequences collected from indoor and outdoor scenes, with a total of 4413 multi-exposure images. Among them, 360 sequences are used for training and 229 sequences are used for testing. Since most images have resolutions ranging from 3000×2000 to 6000×4000 , all images are resized so that the long side is 800 for training and testing.

4.1.2. Implementation Details

The proposed SCAUNet is implemented in PyTorch 2.4.1 and trained on a single NVIDIA GeForce RTX 4090 GPU. The network is trained using 128×128 image patches with a batch size of 8, 8 data-loading workers, and a fixed random seed of 100. Adam optimization with 0.9 and 0.999 is adopted, where the initial learning rate is set to 2×10^{-4} . A cosine annealing restart schedule is employed for 150 K iterations, with the learning rate

gradually decayed to 10^{-6} . Automatic mixed-precision training, gradient clipping, and mixup augmentation are used to stabilize training. For the SICE dataset [35], one image is randomly sampled from each training sequence and then cropped into patches, and this sampling process is repeated 8 times to increase the diversity of training data. The parameters λ_{sc} and λ_{dc} are empirically set to 0.2 and 0.1, respectively. Code is available at <https://github.com/lixiaofang-96/SCAUNet> (accessed on 29 May 2026).

4.1.3. Evaluation Metrics

For images with ground truth, quantitative performance is evaluated using peak signal-to-noise ratio (PSNR) and SSIM [29]. For no-reference datasets, the perception-based image quality evaluator (PIQE) [36] and entropy are adopted. PIQE assesses perceptual quality based on block-wise distortion analysis, while entropy measures the richness of image information. Higher PSNR, SSIM, and entropy values indicate better image quality, whereas a lower PIQE value indicates better perceptual quality. In addition, model complexity is reported to demonstrate the computational efficiency of the proposed one-step unfolding framework.

4.1.4. Baselines

The proposed method is compared with several representative LLIE approaches from different categories. These include the direct mapping method SNR-Net [4]; the STD-based method STANet [12]; Retinex-based methods such as RetinexNet [18], KinD [19], KinD++ [20], URetinex-Net [21], URetinex-Net++ [22], Diff-Retinex++ [25], Retinexformer [16], and Retinexmamba [17]; transfer function-based methods including TBEFN [26] and SAT-Net [27]; and the degradation-aware method CERL [28]. These methods have been trained on the LOL dataset, and then the provided weights are utilized for evaluation.

4.2. Comparison with State-of-the-Art Methods

Table 1 shows that SCAUNet achieves strong quantitative performance on both LOL-v1 and LOL-v2. On LOL-v1 and LOL-v2, the proposed SCAUNet ranks first on both PSNR and SSIM. This result demonstrates that more effective illumination correction and structure preservation can be achieved by the proposed method. Compared with direct mapping, STD-based, Retinex-based, transfer-function-based, and degradation-aware methods, the proposed method gives more favorable results. These results verify the advantage of combining degradation-oriented modeling with the proposed step-size-consistent one-step unfolding framework.

Figures 2 and 3 show the visual comparisons on LOL-v1 and LOL-v2, respectively. In Figure 2, the result of SNR-Net [4] is relatively blurry in local details. STANet [12] does not sufficiently brighten the scene, and noticeable color deviation can be observed. URetinex-Net++ [22] improves overall visibility well, but residual artifacts and color deviation are still visible. Diff-Retinex++ [25] also exhibits slight color deviation, especially in the wooden region. Retinexmamba [17] suffers from residual artifacts and color inconsistency. SATNet [27] also shows slight color deviation. CERL [28] produces visible artifacts, color deviation, and over-smoothed results. By contrast, SCAUNet restores clearer details and preserves colors more faithfully. As a result, the enhanced image is visually closer to the ground truth. In Figure 3, the differences become more evident in the highlighted region. Over-enhancement is visible in STANet [12], URetinex-Net++ [22], Diff-Retinex++ [25], and SATNet [27], where the bright area appears less natural. SNR-Net [4] and Retinexmamba [17] yield relatively better restoration, but residual noise or artifacts can still be observed. CERL [28] again shows noticeable color distortion and over-smoothing. In comparison, SCAUNet restores a more natural brightness level and preserves clearer structures in the highlighted region.

Table 1. The average PSNR and SSIM results on LOL-v1 [18] and LOL-v2 [30] datasets. The best results are in bold and the second best results are underlined.

Type	Methods	Sources	LOL-v1 [18]		LOL-v2 [30]	
			PSNR $\uparrow$	SSIM $\uparrow$	PSNR $\uparrow$	SSIM $\uparrow$
Direct mapping	SNR-Net [4]	CVPR 2022	<u>26.8098</u>	0.8522	27.3678	0.8715
STD-based	STANet [12]	TCSVT 2022	22.0524	0.8195	18.8393	0.7252
Retinex-based	RetinexNet [18]	BMVC 2018	18.8722	0.4251	18.1936	0.4408
	KinD [19]	MM 2019	20.3792	0.8302	23.7833	0.8739
	KinD++ [20]	IJCV 2021	21.8035	0.8357	22.1419	0.8529
	URetinex-Net [21]	CVPR 2022	21.3282	0.8355	21.8353	0.8550
	URetinex-Net++ [22]	TPAMI 2025	23.8261	0.8411	23.9674	<u>0.8777</u>
	Diff-Retinex++ [25]	TPAMI 2025	24.5799	<u>0.8658</u>	23.2443	0.8703
	Retinexformer [16]	ICCV 2023	26.2869	0.8418	27.7071	0.8564
	Retinexmamba [17]	ICONIP 2024	26.2629	0.8400	<u>27.8348</u>	0.8654
Transfer function-based	TBEFN [26]	TMM 2020	17.3510	0.7823	22.4105	0.8547
	SATNet [27]	EAAI 2025	26.1689	0.8631	27.0664	0.8520
Degradation-aware	CERL [28]	TIP 2022	22.6399	0.7500	25.5495	0.7678
	SCAUNet	Ours	27.3195	0.8841	28.5326	0.8972

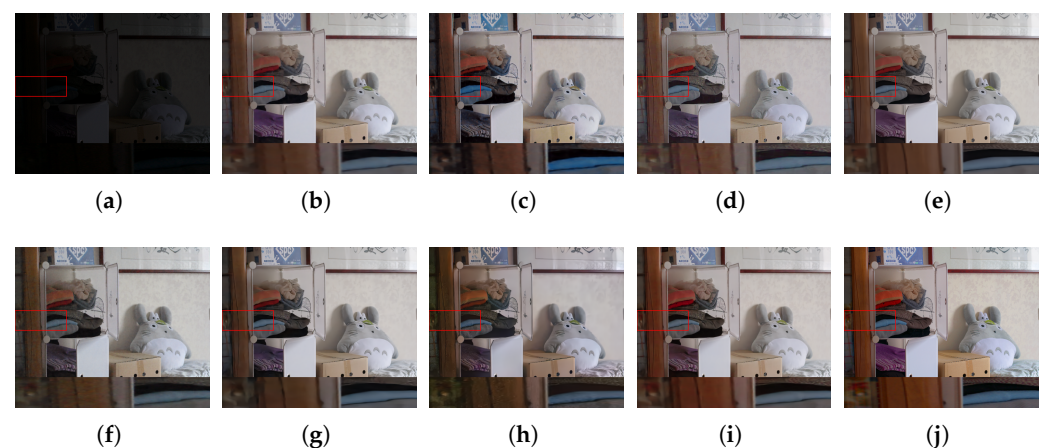
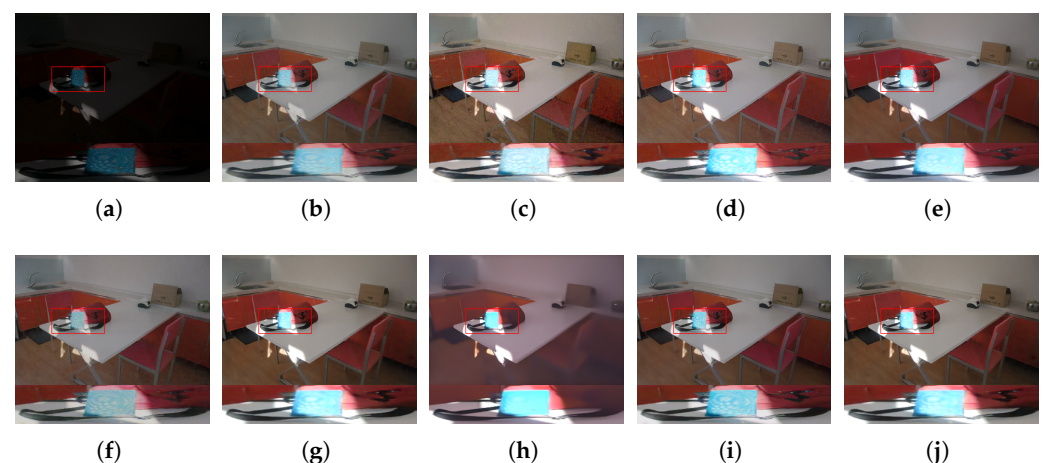
**Figure 2.** Visual comparison of a test image from LOL-v1 [18]. Zoomed-in views are shown in the red boxes. (a) Input. (b) SNR-Net [4]. (c) STANet [12]. (d) URetinex-Net++ [22]. (e) Diff-Retinex++ [25]. (f) Retinexmamba [17]. (g) SATNet [27]. (h) CERL [28]. (i) SCAUNet. (j) Ground Truth.**Figure 3.** Visual comparison of a test image from LOL-v2 [30]. Zoomed-in views are shown in the red boxes. (a) Input. (b) SNR-Net [4]. (c) STANet [12]. (d) URetinex-Net++ [22]. (e) Diff-Retinex++ [25]. (f) Retinexmamba [17]. (g) SATNet [27]. (h) CERL [28]. (i) SCAUNet. (j) Ground Truth.

Table 2 reports the PIQE [36] and entropy results on four no-reference low-light datasets, where the model trained on LOL-v1 [18] is directly applied without any additional fine-tuning. Competitive cross-dataset generalization can be observed. The proposed SCAUNet achieves the best PIQE [36] on DICM [32] and NPE [34], and yields competitive results on LIME [31] and MEF [33]. It also attains competitive entropy values across these datasets. This behavior indicates that the proposed degradation-oriented formulation is not restricted to the LOL-v1 [18] distribution. It also suggests that better robustness can be obtained under unseen real low-light conditions. Figures 4–7 present the visual comparisons on four no-reference datasets. In these figures, several competing methods suffer from artifacts, color deviation, and loss of local details in the highlighted regions. In contrast, SCAUNet produces more balanced illumination, clearer structures, and more natural colors across different scenes. These results are consistent with the no-reference quantitative comparisons. They further show that the proposed method generalizes better across datasets when transferred from LOL-v1 [18] to unseen low-light images without fine-tuning.

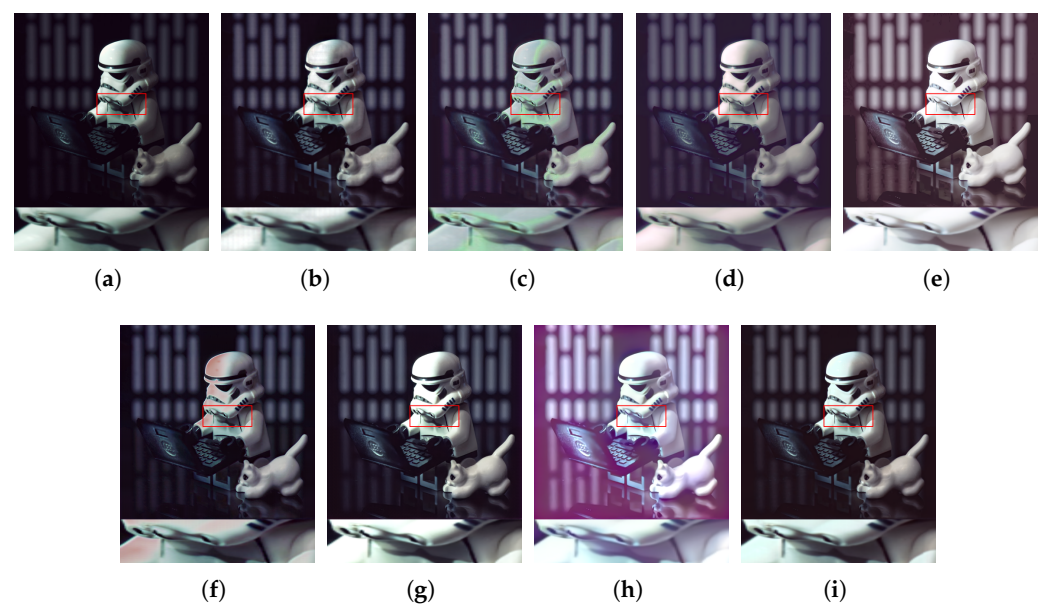


Figure 4. Visual comparison of a test image from LIME [31]. Zoomed-in views are shown in the red boxes. (a) Input. (b) SNR-Net [4]. (c) STANet [12]. (d) URetinex-Net++ [22]. (e) Diff-Retinex++ [25]. (f) Retinexmamba [17]. (g) SATNet [27]. (h) CERL [28]. (i) SCAUNet.

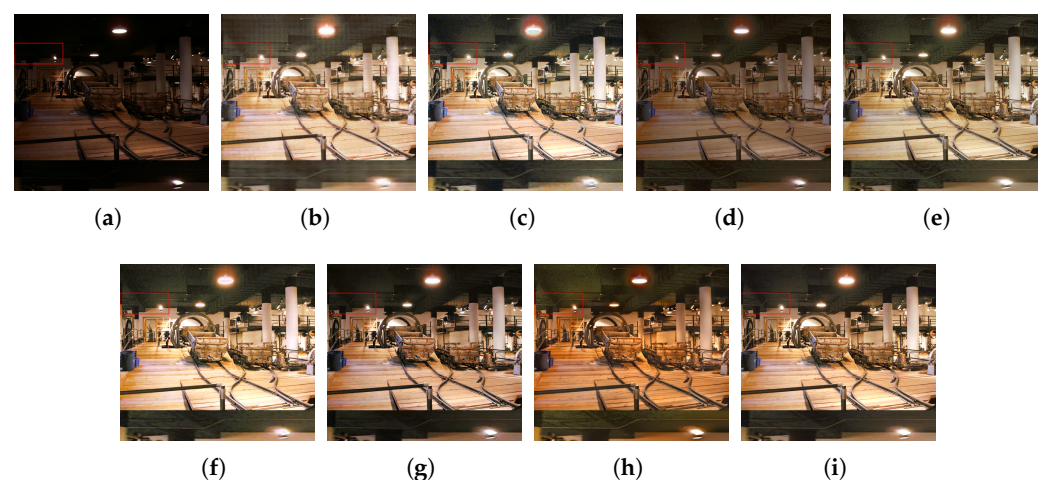


Figure 5. Visual comparison of a test image from DICM [32]. Zoomed-in views are shown in the red boxes. (a) Input. (b) SNR-Net [4]. (c) STANet [12]. (d) URetinex-Net++ [22]. (e) Diff-Retinex++ [25]. (f) Retinexmamba [17]. (g) SATNet [27]. (h) CERL [28]. (i) SCAUNet.

Table 2. The average PIQE and entropy results on no-reference datasets. The best results are in bold and the second best results are underlined.

Methods	LIME [31]		DICM [32]		MEF [33]		NPE [34]	
	PIQE ↓	Entropy ↑	PIQE ↓	Entropy ↑	PIQE ↓	Entropy ↑	PIQE ↓	Entropy ↑
SNR-Net [4]	43.4861	<u>7.2835</u>	<u>28.0874</u>	6.8790	45.6389	7.1673	<u>37.1057</u>	7.1942
STANet [12]	<u>29.3130</u>	7.2290	43.5639	6.6817	58.8040	7.3020	41.4105	7.0023
URetinex-Net++ [22]	32.2512	6.7722	30.5943	7.1092	58.5429	6.9754	47.8964	7.1572
Diff-Retinex++ [25]	28.3898	7.4025	34.7829	7.4092	71.5427	7.4366	44.1923	7.4221
Retinexmamba [17]	51.8407	7.1553	33.1033	7.0509	23.7783	7.1320	44.6784	7.3017
SATNet [27]	51.4835	7.0085	42.2616	7.1980	<u>38.2301</u>	7.1194	39.6575	<u>7.3107</u>
CERL [28]	66.9803	7.2458	28.7348	7.0529	59.6879	7.2194	43.2145	7.2748
SCAUNet	33.5764	7.2585	21.6731	<u>7.3165</u>	39.9987	<u>7.3727</u>	21.8850	7.2673

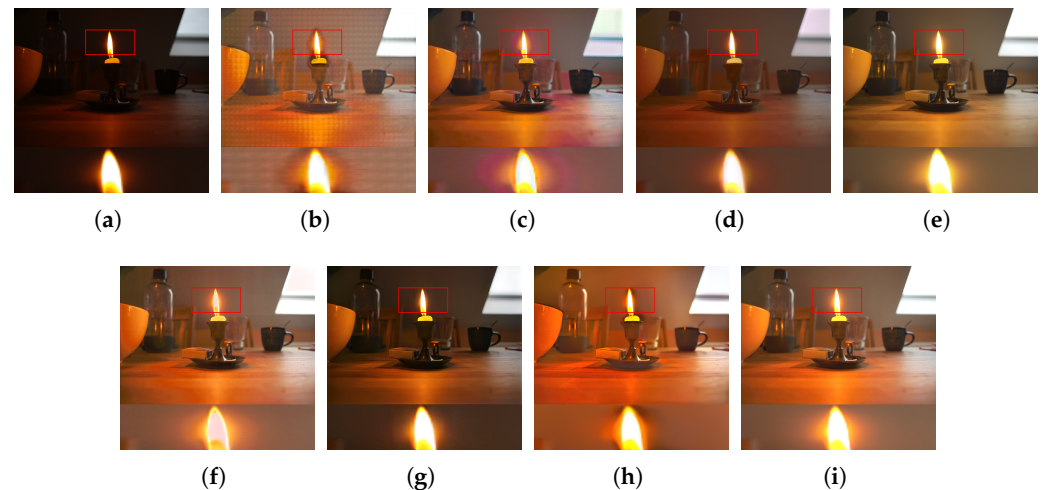


Figure 6. Visual comparison of a test image from MEF [33]. Zoomed-in views are shown in the red boxes. (a) Input. (b) SNR-Net [4]. (c) STANet [12]. (d) URetinex-Net++ [22]. (e) Diff-Retinex++ [25]. (f) Retinexmamba [17]. (g) SATNet [27]. (h) CERL [28]. (i) SCAUNet.

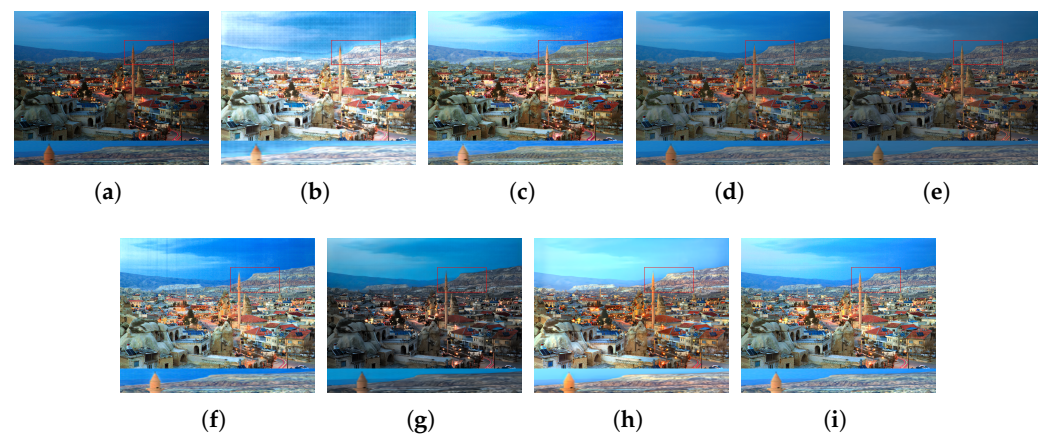


Figure 7. Visual comparison of a test image from NPE [34]. Zoomed-in views are shown in the red boxes. (a) Input. (b) SNR-Net [4]. (c) STANet [12]. (d) URetinex-Net++ [22]. (e) Diff-Retinex++ [25]. (f) Retinexmamba [17]. (g) SATNet [27]. (h) CERL [28]. (i) SCAUNet.

4.3. Applied to Single-Image Exposure Correction

To further evaluate the generality of the proposed framework, it is applied to single-image exposure correction on the SICE dataset [35]. Compared with standard low-light enhancement benchmarks, SICE [35] covers a wider range of exposure conditions. This makes it a more suitable benchmark for evaluating luminance correction. Figure 8 presents the visual results on SICE for the same scene under different input exposure levels. Consistent correction behavior can be observed across all inputs. Under-exposed images are

effectively brightened, while over-exposed images are pulled back to more balanced illumination. Meanwhile, clear structures and natural colors are preserved in the corrected results. These observations suggest that the proposed SCAUNet is not limited to low-light image enhancement, and can be effectively extended to more general single-image exposure correction.



Figure 8. Visual results of single-image exposure correction under different input exposure levels on a test image from SICE [35]. For each exposure image, the corresponding corrected result produced by SCAUNet is shown. The ground-truth image is provided in (m). Zoomed-in views are shown in the red boxes.

5. Discussion

5.1. Ablation Studies

To evaluate the contribution of each component, ablation experiments are conducted on LOL-v1 [18]. As shown in Table 3, the full model achieves the best PSNR and SSIM. When the proposed step-size consistency strategy is removed (w/o SC), both metrics decrease. This confirms that the proposed training strategy is beneficial for stable one-step unfolding. It is noted that the “w/o SC” setting also corresponds to the one-step unfolding case ($iter = 1$). To further examine whether simply increasing the unfolding depth can improve the performance, an $iter = 2$ variant is also tested. As shown in Table 3, directly increasing the iteration number does not lead to further improvement. A possible reason is that, although a larger number of unfolded steps may be beneficial in principle, the deeper unfolded network becomes more difficult to optimize in practice. When the step-size consistency loss $\mathcal{L}_{sc}$ is removed (w/o $\mathcal{L}_{sc}$), the largest PSNR drop is observed, although a relatively competitive SSIM is still retained. This indicates that $\mathcal{L}_{sc}$ plays an important role in improving restoration fidelity. When the dark channel loss $\mathcal{L}_{dc}$ is removed (w/o $\mathcal{L}_{dc}$), the quantitative performance also decreases. In addition, the visual result in Figure 9 becomes less vivid, and the local contrast is weakened. Furthermore, when the initialization of G^0 is replaced by $G^0 = \text{Repeat}\left(\text{AvgPool}\left(\max_{c \in \{R, G, B\}} O_l^{(c)}\right)\right)$ (replace G^0), inferior results are again obtained. This verifies that a physically meaningful initialization is important for reliable state transition. The visual comparisons in Figure 9 further support these

observations. In particular, clearer digits and sharper local structures are preserved by the full model in the highlighted region, while fewer artifacts and more natural brightness are maintained. These results confirm that the proposed step-size consistency constraint, the dark channel loss, and the initialization design all contribute to the final performance.

Table 3. The average PSNR and SSIM results of ablation experiments on LOL-v1 [18]. The best results are in bold.

Methods	LOL-v1	
	PSNR $\uparrow$	SSIM $\uparrow$
w/o SC ($iter = 1$)	26.8269	0.8748
$iter = 2$	26.3307	0.8721
w/o $\mathcal{L}_{sc}$	26.3275	0.8801
w/o $\mathcal{L}_{dc}$	26.7814	0.8747
replace G^0	26.8057	0.8753
Ours	27.3195	0.8841

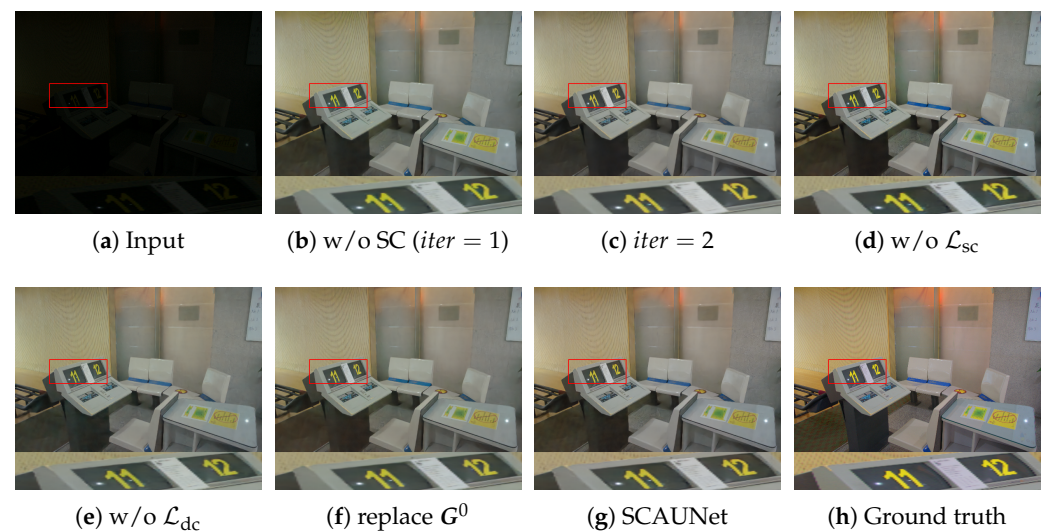


Figure 9. Visual comparison of ablation experiments on a image from LOL-v1 [18]. Zoomed-in views are shown in the red boxes.

To further examine the effect of the proposed step-size consistency constraint, the results produced by one large step and two successive half steps are provided in Figure 10. A clear transition from the intermediate state O_n^{half} to the final restored image O_n^{small} can be observed. Meanwhile, the final result of two half steps O_n^{small} is very close to that produced by one large step O_n^{big} . Similar brightness, structural details, and color appearance are obtained in both cases, especially in the highlighted region. This observation verifies that the proposed step-size consistency constraint effectively regularizes the one-step unfolding process and improves its stability.

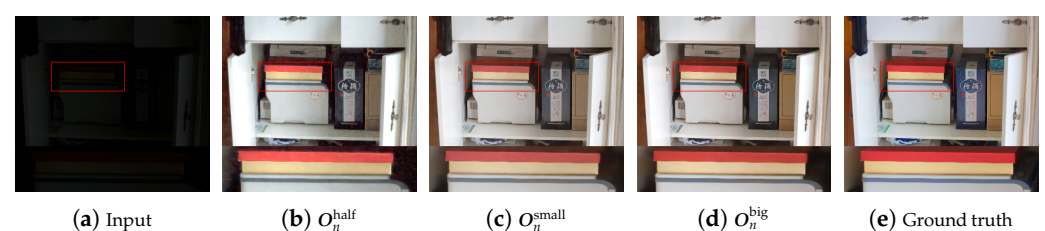


Figure 10. Visual comparison of the results produced by one large step and two successive half steps on a test image from LOL-v1 [18]. Zoomed-in views are shown in the red boxes.

5.2. Parameter Analysis

To investigate the influence of the loss weights, parameter analyses are performed on $\lambda_{sc} = \{0, 0.1, 0.2, 0.5\}$ and $\lambda_{dc} = \{0, 0.01, 0.1, 0.2\}$, and the corresponding results are shown in Figure 11. As λ_{sc} increases from 0 to 0.2, both PSNR and SSIM improve, which verifies the effectiveness of the step-size consistency constraint in stabilizing one-step unfolding. When λ_{sc} is further increased to 0.5, the performance drops slightly, indicating that an excessively large consistency weight may hinder restoration. A similar trend can be observed for λ_{dc} . The best results are obtained at $\lambda_{dc} = 0.1$, while both smaller and larger values lead to inferior performance. This suggests that an appropriate dark channel constraint is beneficial, whereas over-regularization may negatively affect image recovery. Based on these observations, $\lambda_{sc} = 0.2$ and $\lambda_{dc} = 0.1$ are used in the experiments.

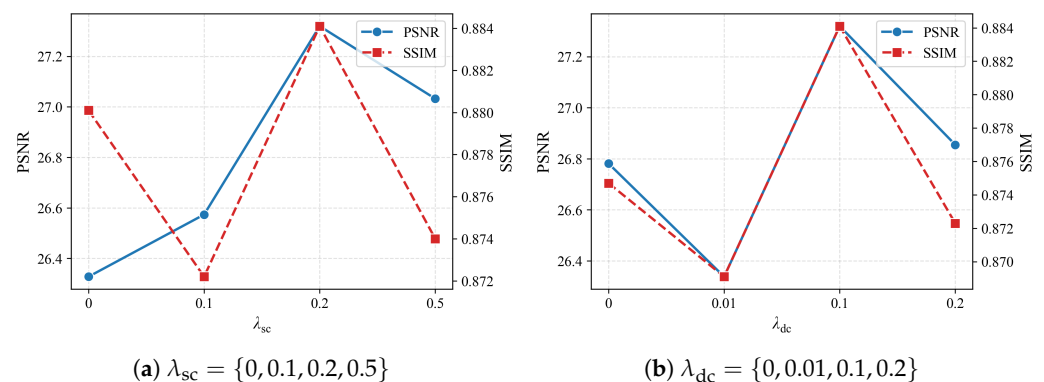


Figure 11. Parameter analysis of the loss weights λ_{sc} and λ_{dc} on LOL-v1 [18].

5.3. Computational Complexity

Table 4 reports the model complexity comparison in terms of floating point operations (FLOPs), the number of trainable parameters on $256 \times 256 \times 3$ inputs, and the average running time on LOL-v1 [18]. SCAUNet maintains a moderate computational cost. Its FLOPs are close to those of Retinexmamba [17] and CERL [28], and remain far below those of Diff-Retinex++ [25] and SATNet [27]. At the same time, the number of trainable parameters is much smaller than that of several strong baselines, including SNR-Net [4], Diff-Retinex++ [25], and SATNet [27]. Moreover, the average running time of SCAUNet remains moderate and is much lower than that of Diff-Retinex++ [25] and SATNet [27]. This indicates that the proposed method achieves a favorable trade-off between enhancement performance and computational cost. The result also supports the efficiency of the proposed one-step unfolding framework, since strong enhancement performance is obtained without introducing excessive model size.

Table 4. Quantitative comparisons of computational complexity in terms of FLOPs (in G), number of trainable parameters (Params) (in M), and the average running time (in s) on LOL-v1 [18].

Methods	FLOPs/G	Params/M	Running Time/s
SNR-Net [4]	24.37	39.12	0.2276
STANet [12]	28.19	0.472	0.4861
URetinex-Net++ [22]	70.83	0.34	0.1515
Diff-Retinex++ [25]	618.01	62.80	23.1389
Retinexmamba [17]	45.98	4.59	0.2931
SATNet [27]	110.68	25.00	1.4820
CERL [28]	44.32	0.67	0.2325
SCAUNet	45.77	5.63	0.3213

5.4. Strengths and Limitations

The proposed method has several advantages. First, a degradation-oriented formulation is adopted, which provides better interpretability than direct black box enhancement. Second, the proposed state-based one-step unfolding framework achieves favorable efficiency while maintaining strong restoration performance. Third, the step-size consistency strategy improves the stability and reliability of one-step unfolding. Nevertheless, several limitations still remain. The current model mainly focuses on luminance degradation and noise corruption, while more complex degradations, such as blur, are not explicitly considered. In addition, the present study mainly evaluates restoration quality at the image level. Its influence on perceptual quality assessment [37] and the robustness of downstream vision systems, especially under adversarial attacks [38,39] and other challenging perturbations, has not yet been investigated. This issue is important because image enhancement may affect both perceptual quality evaluation and the feature distribution of subsequent models. More general degradation modeling, perceptual quality assessment, and robustness-oriented evaluation will be considered in future work.

6. Conclusions

In this paper, a degradation-oriented framework has been proposed for low-light image enhancement. A low-light image is modeled as the element-wise product of a normal-light image and a luminance degradation operator, together with additive noise. Based on this formulation, low-light enhancement is converted into the joint estimation of the target normal-light image and the degradation operator. To solve the resulting problem efficiently, a state-based one-step ADMM solver has been developed, where the main variables are updated by learnable solvers and the remaining variables are updated analytically. In addition, a step-size consistency constraint has been introduced to improve the stability and reliability of the proposed one-step unfolding network SCAUNet.

Extensive experiments on LOL-v1 and LOL-v2 have demonstrated the effectiveness of the proposed SCAUNet. Strong generalization ability has also been verified on several no-reference low-light datasets. Furthermore, the proposed framework has been extended to single-image exposure correction on SICE, where promising visual results have been obtained under different exposure conditions. These results have shown that the proposed method can achieve a favorable balance among illumination correction, detail preservation, artifact suppression, and computational efficiency.

In future work, a more general degradation model will be considered to further improve robustness under complex real imaging conditions. In addition, more flexible unfolding strategies and more unified learning schemes for low-light enhancement and single-image exposure correction will be further investigated.

Author Contributions: Conceptualization, X.L. and C.F.; methodology, X.L.; software, X.L. and H.T.; validation, X.L.; formal analysis, X.L.; data curation, H.T.; writing—original draft preparation, X.L. and H.T.; writing—review and editing, C.F.; visualization, X.L. and H.T.; supervision, C.F.; project administration, C.F.; funding acquisition, C.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by the funding of Xianyang Normal University, grant number 2025ZXY07, and in part by the Special Scientific Research Program of Shaanxi Provincial Education Department, grant number 24JK0704.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhao, C.; Yue, W.; Wang, Y.; Wang, J.; Luo, S.; Chen, H.; Wang, Y. Low-Light Image Enhancement Integrating Retinex-Inspired Extended Decomposition with a Plug-and-Play Framework. *Mathematics* **2024**, *12*, 4025. [\[CrossRef\]](#)
2. He, F.; Liu, B.; Dong, Z.; Li, J.; Luo, J.; Zhao, D. An Energy-Efficient Zero-Shot AI-ISP for Real-Time Low-Light Enhancement with Intelligent Vehicles. *Mathematics* **2026**, *14*, 1324. [\[CrossRef\]](#)
3. Son, D.M.; Lee, S.H. Low-Light Image Enhancement for Driving Condition Recognition Through Multi-Band Images Fusion and Translation. *Mathematics* **2025**, *13*, 1418. [\[CrossRef\]](#)
4. Xu, X.; Wang, R.; Fu, C.W.; Jia, J. SNR-Aware Low-Light Image Enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; IEEE: New York, NY, USA, 2022; pp. 17714–17724. [\[CrossRef\]](#)
5. Wang, Y.; Wan, R.; Yang, W.; Li, H.; Chau, L.P.; Kot, A. Low-Light Image Enhancement with Normalizing Flow. In *Proceedings of the AAAI Conference on Artificial Intelligence*; Association for the Advancement of Artificial Intelligence: Palo Alto, CA, USA, 2022; Volume 36, pp. 2604–2612. [\[CrossRef\]](#)
6. Xu, L.; Hu, C.; Hu, Y.; Jing, X.; Cai, Z.; Lu, X. UPT-Flow: Multi-Scale Transformer-Guided Normalizing Flow for Low-Light Image Enhancement. *Pattern Recognit.* **2025**, *158*, 111076. [\[CrossRef\]](#)
7. Weng, J.; Yan, Z.; Tai, Y.; Qian, J.; Yang, J.; Li, J. MambaLLIE: Implicit Retinex-aware Low Light Enhancement with Global-then-local State Space. In *Proceedings of the 38th International Conference on Neural Information Processing Systems*; NeurIPS: San Diego CA, USA, 2024; pp. 27440–27462. [\[CrossRef\]](#)
8. Jiang, H.; Luo, A.; Liu, X.; Han, S.; Liu, S. LightenDiffusion: Unsupervised Low-Light Image Enhancement with Latent-Retinex Diffusion Models. In *Proceedings of the European Conference on Computer Vision*; Springer: Cham, Switzerland, 2024; pp. 161–179. [\[CrossRef\]](#)
9. Yang, S.; Zhang, X.; Wang, Y.; Yu, J.; Wang, Y.; Zhang, J. DiffLLE: Diffusion-Based Domain Calibration for Weak Supervised Low-Light Image Enhancement. *Int. J. Comput. Vis.* **2025**, *133*, 2527–2546. [\[CrossRef\]](#)
10. Xu, K.; Yang, X.; Yin, B.; Lau, R.W. Learning to Restore Low-Light Images via Decomposition-and-Enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; IEEE: New York, NY, USA, 2020; pp. 2281–2290. [\[CrossRef\]](#)
11. Zheng, C.; Shi, D.; Shi, W. Adaptive Unfolding Total Variation Network for Low-Light Image Enhancement. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*; IEEE: New York, NY, USA, 2021; pp. 4439–4448. [\[CrossRef\]](#)
12. Xu, K.; Chen, H.; Xu, C.; Jin, Y.; Zhu, C. Structure-Texture Aware Network for Low-Light Image Enhancement. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 4983–4996. [\[CrossRef\]](#)
13. Jiang, H.; Luo, A.; Fan, H.; Han, S.; Liu, S. Low-Light Image Enhancement with Wavelet-Based Diffusion Models. *ACM Trans. Graph.* **2023**, *42*, 1–14. [\[CrossRef\]](#)
14. Wang, R.; Zhang, Q.; Fu, C.W.; Shen, X.; Zheng, W.S.; Jia, J. Underexposed Photo Enhancement Using Deep Illumination Estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; IEEE: New York, NY, USA, 2019; pp. 6849–6857. [\[CrossRef\]](#)
15. Liu, R.; Ma, L.; Zhang, J.; Fan, X.; Luo, Z. Retinex-Inspired Unrolling with Cooperative Prior Architecture Search for Low-Light Image Enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; IEEE: New York, NY, USA, 2021; pp. 10561–10570. [\[CrossRef\]](#)
16. Cai, Y.; Bian, H.; Lin, J.; Wang, H.; Timofte, R.; Zhang, Y. Retinexformer: One-stage Retinex-based Transformer for Low-Light Image Enhancement. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*; IEEE: New York, NY, USA, 2023; pp. 12504–12513. [\[CrossRef\]](#)
17. Bai, J.; Yin, Y.; He, Q.; Li, Y.; Zhang, X. Retinexmamba: Retinex-Based Mamba for Low-Light Image Enhancement. In *Proceedings of the International Conference on Neural Information Processing*; Springer: Singapore, 2025; pp. 427–442. [\[CrossRef\]](#)
18. Wei, C.; Wang, W.; Yang, W.; Liu, J. Deep Retinex Decomposition for Low-Light Enhancement. In *Proceedings of the British Machine Vision Conference*; British Machine Vision Association: Durham, UK, 2018; pp. 1–12. [\[CrossRef\]](#)
19. Zhang, Y.; Zhang, J.; Guo, X. Kindling the Darkness: A Practical Low-Light Image Enhancer. In *Proceedings of the 27th ACM International Conference on Multimedia*; Association for Computing Machinery: New York, NY, USA, 2019; pp. 1632–1640. [\[CrossRef\]](#)
20. Zhang, Y.; Guo, X.; Ma, J.; Liu, W.; Zhang, J. Beyond Brightening Low-Light Images. *Int. J. Comput. Vis.* **2021**, *129*, 1013–1037. [\[CrossRef\]](#)
21. Wu, W.; Weng, J.; Zhang, P.; Wang, X.; Yang, W.; Jiang, J. URetinex-Net: Retinex-Based Deep Unfolding Network for Low-Light Image Enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; IEEE: New York, NY, USA, 2022; pp. 5901–5910. [\[CrossRef\]](#)
22. Wu, W.; Weng, J.; Zhang, P.; Wang, X.; Yang, W.; Jiang, J. Interpretable Optimization-Inspired Unfolding Network for Low-Light Image Enhancement. *IEEE Trans. Pattern Anal. Mach. Intell.* **2025**, *47*, 2545–2562. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Liu, X.; Xie, Q.; Zhao, Q.; Wang, H.; Meng, D. Low-Light Image Enhancement by Retinex-Based Algorithm Unrolling and Adjustment. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, *35*, 15758–15771. [\[CrossRef\]](#)

24. Zheng, N.; Zhou, M.; Dong, Y.; Rui, X.; Huang, J.; Li, C.; Zhao, F. Empowering Low-Light Image Enhancer through Customized Learnable Priors. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*; IEEE: New York, NY, USA, 2023; pp. 12525–12535. [[CrossRef](#)]
25. Yi, X.; Xu, H.; Zhang, H.; Tang, L.; Ma, J. Diff-Retinex++: Retinex-Driven Reinforced Diffusion Model for Low-Light Image Enhancement. *IEEE Trans. Pattern Anal. Mach. Intell.* **2025**, *47*, 6823–6841. [[CrossRef](#)] [[PubMed](#)]
26. Lu, K.; Zhang, L. TBEFN: A Two-Branch Exposure-Fusion Network for Low-Light Image Enhancement. *IEEE Trans. Multimed.* **2020**, *23*, 4093–4105. [[CrossRef](#)]
27. Li, X.; Wang, W.; Han, Y.; Feng, X. Structure Aware Transfer Function Network for Low Light Image Enhancement. *Eng. Appl. Artif. Intell.* **2025**, *141*, 109749. [[CrossRef](#)]
28. Chen, Z.; Jiang, Y.; Liu, D.; Wang, Z. CERL: A Unified Optimization Framework for Light Enhancement with Realistic Noise. *IEEE Trans. Image Process.* **2022**, *31*, 4162–4172. [[CrossRef](#)]
29. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)]
30. Yang, W.; Wang, W.; Huang, H.; Wang, S.; Liu, J. Sparse Gradient Regularized Deep Retinex Network for Robust Low-Light Image Enhancement. *IEEE Trans. Image Process.* **2021**, *30*, 2072–2086. [[CrossRef](#)] [[PubMed](#)]
31. Guo, X.; Li, Y.; Ling, H. LIME: Low-light image enhancement via illumination map estimation. *IEEE Trans. Image Process.* **2016**, *26*, 982–993. [[CrossRef](#)]
32. Lee, C.; Lee, C.; Kim, C.S. Contrast Enhancement based on Layered Difference Representation of 2D Histograms. *IEEE Trans. Image Process.* **2013**, *22*, 5372–5384. [[CrossRef](#)]
33. Ma, K.; Zeng, K.; Wang, Z. Perceptual Quality Assessment for Multi-Exposure Image Fusion. *IEEE Trans. Image Process.* **2015**, *24*, 3345–3356. [[CrossRef](#)]
34. Wang, S.; Zheng, J.; Hu, H.M.; Li, B. Naturalness Preserved Enhancement Algorithm for Non-Uniform Illumination Images. *IEEE Trans. Image Process.* **2013**, *22*, 3538–3548. [[CrossRef](#)]
35. Cai, J.; Gu, S.; Zhang, L. Learning a Deep Single Image Contrast Enhancer from Multi-Exposure Images. *IEEE Trans. Image Process.* **2018**, *27*, 2049–2062. [[CrossRef](#)] [[PubMed](#)]
36. N, V.; D, P.; Bh, M.C.; Channappayya, S.S.; Medasani, S.S. Blind Image Quality Evaluation Using Perception based Features. In *Proceedings of the National Conference on Communications*; IEEE: New York, NY, USA, 2015; pp. 1–6. [[CrossRef](#)]
37. Lin, J.; Cui, Y.; Fang, C.; Pan, B.; Pan, C.; Jiang, G.; Zhang, S.; Ma, S.; Tian, Q. Multi-Modal Cross-Attention-Guided Network for Audio-Visual Quality Evaluation via Visual Saliency and Mel-Spectrum Features. *IEEE Trans. Circuits Syst. Video Technol.* **2026**, *36*, 6783–6798. [[CrossRef](#)]
38. Dang, Q.; Ruan, J.; Zhan, T.; Gong, M.; He, X. LADA: Latent-Space Adversarial Diffusion Attack in Remote Sensing. *IEEE Trans. Geosci. Remote Sens.* **2025**, *63*, 5631616. [[CrossRef](#)]
39. Dang, Q.; Zhan, T.; Gong, M.; He, X. Boosting Adversarial Transferability by Batchwise Amplitude Spectrum Normalization. *IEEE Trans. Geosci. Remote Sens.* **2025**, *63*, 5608114. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.