Computers, Materials & Continua

Tech Science Press

DOI: 10.32604/cmc.2024.059000

ARTICLE

Unsupervised Low-Light Image Enhancement Based on Explicit Denoising and Knowledge Distillation

Wenkai Zhang[1,2], Hao Zhang[1,2], Xianming Liu[1], Xiaoyu Guo[1,2], Xinzhe Wang[1] and Shuiwang Li[1,2,*]

[1]College of Computer Science and Engineering, Guilin University of Technology, Guilin, 541006, China

[2]Guangxi Key Laboratory of Embedded Technology and Intelligent System, Guilin University of Technology, Guilin, 541004, China

*Corresponding Author: Shuiwang Li. Email: lishuiwang0721@163.com

Received: 25 September 2024 Accepted: 14 November 2024 Published: 17 February 2025

ABSTRACT

Under low-illumination conditions, the quality of image signals deteriorates significantly, typically characterized by a peak signal-to-noise ratio (PSNR) below 10 dB, which severely limits the usability of the images. Supervised methods, which utilize paired high-low light images as training sets, can enhance the PSNR to around 20 dB, significantly improving image quality. However, such data is challenging to obtain. In recent years, unsupervised low-light image enhancement (LIE) methods based on the Retinex framework have been proposed, but they generally lag behind supervised methods by 5–10 dB in performance. In this paper, we introduce the Denoising-Distilled Retine (DDR) method, an unsupervised approach that integrates denoising priors into a Retinex-based training framework. By explicitly incorporating denoising, the DDR method effectively addresses the challenges of noise and artifacts in low-light images, thereby enhancing the performance of the Retinex framework. The model achieved a PSNR of 19.82 dB on the LOL dataset, which is comparable to the performance of supervised methods. Furthermore, by applying knowledge distillation, the DDR method optimizes the model for real-time processing of low-light images, achieving a processing speed of 199.7 fps without incurring additional computational costs. While the DDR method has demonstrated superior performance in terms of image quality and processing speed, there is still room for improvement in terms of robustness across different color spaces and under highly resource-constrained conditions. Future research will focus on enhancing the model's generalizability and adaptability to address these challenges. Our rigorous testing on public datasets further substantiates the DDR method's state-of-the-art performance in both image quality and processing speed.

KEYWORDS

Deep learning; low-light image enhancement; real-time processing; knowledge distillation

1 Introduction

Low-light images are commonly encountered in everyday photography and autonomous driving scenarios [1]. In nocturnal or low-light environments, the quality of captured images tends to degrade significantly compared to customary conditions, with the primary characteristics being excessive darkness, reduced resolution, and increased noise. This typically manifests as a PSNR below 10 dB and a Structural Similarity Index (SSIM) below 0.5 [2,3].

Low-light image enhancement (LIE) focuses on improving images captured in dim conditions to make them resemble scenes taken in ordinary daylight, making it an essential area of image processing [4–8]. The main objective is to brighten low-light images, revealing more information that is easier for both human observers and machine algorithms to process and analyze [9–11]. LIE techniques have been widely applied in fields such as aerospace, road recognition, biomedicine, disaster relief, and rescue operations [12,13]. For instance, using low-light enhancement technology to enhance medical images facilitates doctors' precise diagnosis of lesion areas; applying LIE to video surveillance solves the problem of complex object recognition in low-light conditions [14].

Images captured in low-light settings are frequently subject to a variety of distortions, including sensor noise, limited visibility, and low contrast [6]. These issues make low-light images unsuitable for effective information sharing, as they hinder both human visual perception and downstream computer vision applications [15]. Over the past several decades, considerable research efforts have been directed toward the development of LIE algorithms, aiming to rectify contrast, uncover textures, and eliminate sensor noise [4,14]. End-to-end low illumination graphs such as the Low Light Net (LLNet) [16] and the multi-branch low light enhancement network (MBLLEN) [17] image enhancement work have demonstrated the possibility of using neural networks to improve the quality of low-illumination images. Researchers have also observed that the Retinex model performs well in traditional LIE and image-defogging tasks. This has led to the development of methods that leverage neural networks to estimate the illumination and reflection components within the Retinex framework. Notable examples include Retinex-Net [2] and LightenNet [18], which have demonstrated strong enhancement capabilities and adaptability in low-light scenarios. To enhance the generalizability of neural networks for LIE, many researchers have developed and collected specialized low-light datasets, such as the See-in-the-Dark (SID) [19] and Low-Light (LOL) [2]. These datasets have become essential for training and evaluating LIE models. In recent years, there has been substantial progress in developing unsupervised LIE methods that do not rely on labeled data, overcoming the limitations of supervised approaches that require extensive annotations. Unsupervised techniques in LIE have gained significant attention due to their ability to improve image quality without the need for labeled datasets. These methods are particularly valuable when collecting annotated data is costly or impractical. One prominent example is Enlighten Generative Adversarial Networks (EnlightenGAN) [10], which enhances images using a GAN-based approach that does not require paired training samples. By leveraging adversarial learning, EnlightenGAN is able to generate high-quality enhanced images even in the absence of ground truth references, making it an effective unsupervised solution. In addition to GAN-based methods, zero-shot learning approaches like Zero-reference Deep Curve Estimation (Zero-DCE) [8] and Paired Low-Light Instances Enhancer (PairLIE) [5] have emerged as powerful tools in unsupervised LIE. These methods stand out for their minimal data requirements and quick adaptability. Zero-DCE, for instance, formulates image enhancement as a curve estimation problem, allowing it to learn effective enhancement mappings directly from the input data without requiring any reference images. Similarly, PairLIE leverages instance-level learning to perform effective low-light enhancement with low data costs, providing an efficient alternative for scenarios where large datasets are not available.

Despite their advantages, unsupervised LIE methods still need to work on balancing brightness enhancement with noise amplification. As images are brightened, noise is often amplified, degrading image clarity and quality. This problem is particularly acute in unsupervised settings where annotated data is not available for fine-tuning the model's response to noise. In this work, we introduce an unsupervised framework specifically designed to address the challenges of LIE by incorporating an explicit denoising subnetwork. Unlike traditional approaches, our framework is guided by a pre-trained denoising model, which provides valuable prior knowledge during the training process.

The design of this subnetwork serves two primary purposes. First, the denoising subnetwork is intentionally lightweight, significantly reducing computational overhead compared to the pre-trained model it learns from. This ensures that our framework remains efficient and suitable for real-time applications, which is crucial in resource-constrained environments such as embedded systems and mobile devices. Second, by integrating the denoising subnetwork with the rest of the architecture, we enable end-to-end training that harmonizes both denoising and image brightening. This balanced approach allows the model to simultaneously enhance image brightness and clarity while effectively mitigating noise, resulting in higher-quality outputs. The end-to-end nature of the training also ensures that the denoising process adapts dynamically to the specific needs of LIE, improving overall performance. In addition, we further enhance the efficiency of our proposed method through the use of knowledge distillation. Knowledge distillation is a technique that transfers knowledge from a large, well-trained "teacher" model to a smaller, more compact "student" model. By incorporating this approach, we can preserve high performance while substantially reducing the model's size and computational complexity, resulting in a more compact and efficient model that is better suited for real-time applications. Extensive experiments have validated the effectiveness of our approach and show that it achieves state-of-art-performance. The contributions of our work can be outlined as follows:

- Our primary contribution is the integration of a pre-trained denoising model into a Retinex-based framework for LIE. This approach significantly reduces noise and artifacts, tackling a key challenge in unsupervised LIE.
- Additionally, we employ knowledge distillation to develop a compact model suitable for resource-constrained environments, achieving both high image quality and efficient real-time processing.
- Comprehensive experiments show that our method narrows the gap between unsupervised and supervised LIE techniques, outperforming existing unsupervised approaches in image quality and speed.

2 Related Work

2.1 Conventional Methods

Conventional LIE techniques are essential for improving image clarity in suboptimal lighting, and they span Histogram-based approaches and Retinex-based models. Histogram-based approaches extend the dynamic range to enhance brightness. For instance, Park et al. [20] segmented the histogram's range, adjusting gray levels based on the area ratio, while Lee et al. [21] used a hierarchical representation to heighten inter-pixel gray level contrasts. Retinex-based methods address low-light issues by separating images into reflectance and illumination components. Enhanced images are created either by using reflectance directly or adjusting illumination and recombining it. Guo et al. [22] estimated illumination by taking maximum RGB values and refining them with structural priors, and Li et al. [23] integrated a noise map to improve Retinex outcomes. The texture-aware Retinex model by Xu et al. [3] optimizes through iterative steps, while Hao et al. [7] proposed a semi-decoupled approach based on Retinex, furthering the robustness of conventional models.

2.2 Learning-Based Methods

Learning methods for LIE often necessitate paired datasets comprising both low-light and well-lit images. Lore et al. [16] crafted a multi-layered sparse denoising autoencoder for LIE, training

their model using artificially generated image pairs. Wei et al. [2] pioneered the creation of a real-world dataset comprising matched low-light and normal-light images, which they employed to train an end-to-end network in a supervised learning framework. Leveraging this dataset, the researchers further developed a fully convolutional neural network tailored for the enhancement of low-light images. Wu et al. [24] introduced a novel deep unfolding network inspired by Retinex theory, aimed at enhancing the network's adaptability and computational efficiency. Xu et al. [25] integrated a signal-to-noise ratio (SNR)–aware transformer with a convolutional neural network to achieve better LIE. Lastly, Zhang et al. [26] devised a network focused on color consistency, aiming to reduce color discrepancies between their respective ground truth images and enhanced images.

Recently, advancements in networks for unsupervised learning have targeted reducing reliance on reference images. For instance, Guo et al. [8] have introduced a LIE method that does not require references, with their network fine-tuned through non-reference loss functions. Jiang et al. [10] have presented a LIE method that leverages generative adversarial networks and data without pairing. Liu et al. [27] combined unfolding methods with strategic prior architecture search in a compact LIE method. Fu et al. [5] introduced an unsupervised model called PairLIE, which used paired low-light images to learn adaptive priors based on. RetinexFormer, Sharif et al. [9] introduce Transformer architectures into LIE, further enhancing the model's expressive power.

In the domain of image classification techniques, recent contributions have innovatively tackled issues related to low-light conditions. Yang et al. [28] proposed an implicit neural representation for cooperative low-light image enhancement, which can potentially have a positive impact on downstream tasks such as image classification by improving image quality under low-light conditions. Additionally, Hashmi et al. [29] focused on enhancing hierarchical features for object detection and beyond under low-light vision, which can boost the performance of classification and detection networks in such challenging lighting conditions.

While supervised models achieve superior quality through verified annotations, unsupervised approaches generally require minimal data preprocessing and simplifying deployment but often face challenges in detail fidelity due to intrinsic noise. Here, a denoising subnetwork bolsters robustness against noise interference. Knowledge distillation techniques complement these methods by speeding up processing while retaining accuracy, and our approach employs these techniques to ensure the system balances efficiency with quality output.

2.3 Knowledge Distillation

Knowledge distillation, initially proposed by Hinton et al. [30], effectively compresses model sizes by transferring learned knowledge from a teacher network to a student network. Chen et al. [31] developed rapid training techniques by transferring function-preserving transformations, while Zhang et al. [32] used self-distillation for knowledge transfer within a single network. Attention map-based methods by Zagoruyko et al. [33] used attentional cues for enhanced knowledge transfer. Liu et al. [34] explored structured distillation through similarity maps and adversarial techniques. Distillation techniques have shown promise in single-image super-resolution, as shown by Gao et al. [35], He et al. [36], and Lee et al. [11]. However, their application for low-light enhancement is scarce. Notably, Li et al. [37] have explored distillation for low-light image tasks, which informs the foundation of our approach.

3 Proposed Method

We begin by introducing the Retinex theory, which forms the foundation of our unsupervised LIE model. Next, we offer a detailed explanation of the proposed network architecture. Finally, we outline the workflow and corresponding loss functions. The subsequent subsections will delve into the specifics of each of these components.

3.1 Unsupervised LIE Model Based on Retinex

In accordance with Retinex theory, a low-light image can be broken down into an illumination component L and a reflectance component R.

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

where $\circ$ represents element-wise multiplication, L represents the light intensity of objects, which should be piecewise continuous and devoid of texture. R represents the physical properties of objects, which should encompass the texture and details visible in the image. The general approach to Retinex decomposition involves minimizing the following energy function:

$$\underset{L,R}{\operatorname{argmin}} \|L \circ R - I\|_2 + \lambda_R f_R(R) + \lambda_L f_L(L), \tag{2}$$

where f_R and f_L are the prior constraints for R and L, respectively. λ_R and λ_L represent the weights. $\|L \circ R - I\|_2$ is the data fidelity term that measures the difference between the input and the reconstructed image. Fu et al. [5] proposed decomposing a pair of images to incorporate additional information and constraints for unsupervised LIE learning. We adopt this approach to build our unsupervised framework. Mathematically, the decomposition of paired low-light images can be formulated as follows:

$$\begin{cases} I_1 = L_1 \circ R \\ I_2 = L_2 \circ R, \end{cases} \tag{3}$$

where I_1 and I_2 constitute a pair of low-light images that share the same reflectance component R. L_1 and L_2 represent different light intensities. Since I_1 and I_2 are a pair of low-light images lacking prior knowledge from ground truth, the proposed DDR model operates in an unsupervised manner. The reflectance and illumination components obtained through the Retinex Decomposition Module provide essential inputs for the upcoming L-Net, R-Net, and F-Net network architectures, supporting specific low-light enhancement processing.

3.2 Network Architecture

In this paper, we advocate a LIE model built upon Retinex theory. Initially, all input low-light images are processed through a Retinex Decomposition Module (RDM) to separate the illumination and reflectance components within the image accurately. Our training procedure is divided into two phases: In the first phase, we focus on training the teacher model, which takes a pair of low-light images as input. The model utilizes information decomposed by the RDM, employs a denoising module for efficient noise reduction and feature extraction, and is optimized with a customized loss function to learn effective denoising and feature enhancement. In the second phase, the student model is trained under the direct guidance of the teacher model. The student model is designed to improve real-time processing capabilities while preserving effective LIE performance. Trained on single images, it employs knowledge distillation techniques to transfer the denoising and feature enhancement expertise from the teacher model. This approach allows the student model to significantly reduce computational

complexity and avoid complex denoising modules while still ensuring real-time solid performance. The subsequent sections of this paper will detail the specific implementation details and testing procedures of this method.

The RDM, as illustrated in Fig. 1, is a critical component of our method, which harnesses the power of three specialized networks: L-Net, R-Net, and F-Net. L-Net is dedicated to estimating the illumination component, while R-Net focuses on the reflectance component of an image. F-Net plays a pivotal role in enhancing the process by removing spurious features from the original image. Each network is streamlined to consist of only five convolutional layers, with the first four layers equipped with the ReLU activation function. The networks culminate in a sigmoid layer that ensures the output values are constrained within the [0, 1] range. In alignment with Retinex theory, L-Net produces a single output channel for the standard illumination across all color channels, whereas R-Net generates three channels to capture the distinct reflectance details of each color. The integration of F-Net boosts the precision of our decomposition technique, particularly for low-light images, resulting in a more refined and accurate separation of the illumination and reflectance components. During the training phase of the Teacher model, we initiate the process by feeding a pair of original low-light images, I_1 and I_2, into the F-Net. This results in the generation of optimized versions, i_1 and i_2. Subsequently, we estimate the potential illuminations (L_{t_1} and L_{t_2}) and reflectances (R_{t_1} and R_{t_2}) using the L-Net and R-Net, respectively. The enhanced images are then computed using Eq. (4) and processed through the Prior Denoising Model (PD-net) and the Denoising Subnetwork (D-net).

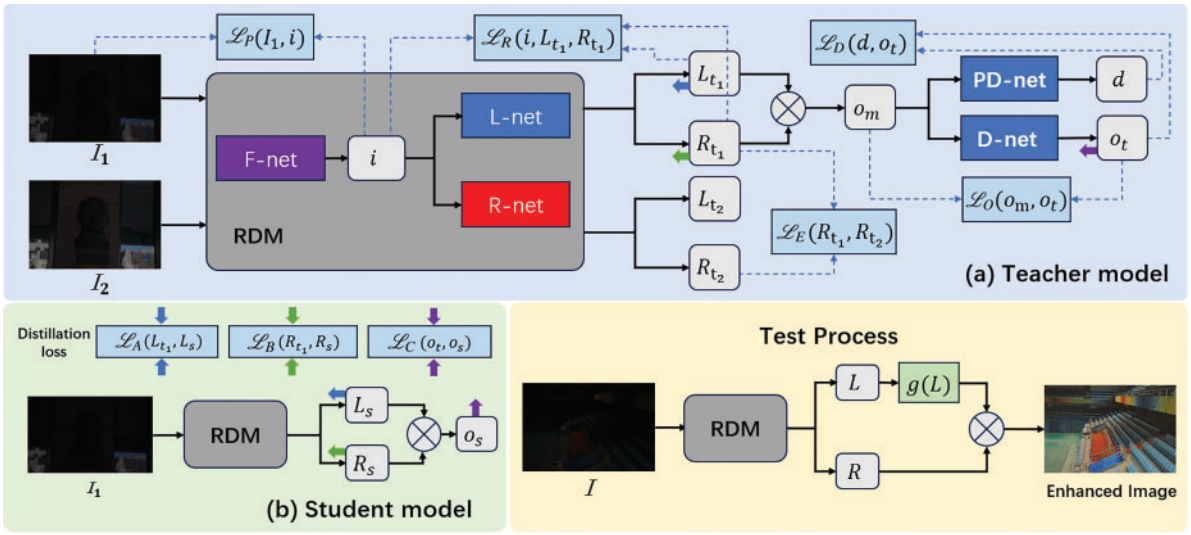

Figure 1: The diagram of the unsupervised LIE model based on Retinex presented in this paper can be described in two phases: Teacher and student. (a) In the first stage, the Teacher model is trained, in which pairs of low-brightness images are used for input, and then multi-scale features of the image mapping after noise reduction are learned to improve the model's capability in reducing noise. (b) In the second stage, distillation learning is carried out using the Student model. In the distillation learning process, the Teacher model only deduces without updating parameters, and the Student model obtains fixed knowledge from the teacher model in each training cycle

For the PD-net, we employ a Self-Guided Neural Network (SGN) [6] that utilizes a distinctive shuffling mechanism to generate multi-scale images. This method effectively leverages information

from low-resolution branches during the high-resolution feature extraction phase, enabling cross-scale information fusion and significantly improving image denoising quality. However, the SGN's complexity and high computational demands present challenges to achieving real-time performance. To overcome this, we develop the D-net, a streamlined shallow neural network that balances denoising effectiveness with computational efficiency, which is crucial for real-time applications. After processing through PD-net and D-net, we obtain the prior denoised feature vector d and the output enhanced image feature o_t, respectively. We boost the denoising performance using the denoising loss $\mathscr{L}_D$ and the Projection Loss $\mathscr{L}_O$. The $\mathscr{L}_D$ measures the alignment between the PD-net's output and the enhanced image feature, while the $\mathscr{L}_O$ evaluates the discrepancy between the original and target image features. By optimizing these loss functions together, we aim to refine the denoising process while efficiently managing computational resources. For the Student model, the training phase requires only a single original low-light image I as input into the RDM to estimate the potential illumination L and reflectance R. To facilitate knowledge distillation from the Teacher to the Student model, we introduce three loss functions: $\mathscr{L}_{Di}$. This structured approach to training ensures that the Student model inherits the denoising and enhancement capabilities of the Teacher model while maintaining a smaller footprint appropriate for real-time applications in resource-constrained environments. The upcoming section focused on loss design will provide a thorough explanation of the design and implementation of these loss functions.

During testing, a low-light image is given, and the final enhanced image is calculated using Eq. (4) by applying F-Net, R-net, and L-net sequentially:

$$I_{en} = R \circ g(L) = R \circ L^{\lambda}, \tag{4}$$

where I_{en} represents the enhanced image, and λ is the illumination correction coefficient.

To achieve optimal denoising performance for PD-net and D-net, we detail specific loss function designs in Section 3.3 to enhance both denoising and feature extraction quality.

3.3 Training Losses

In our method, the loss function is composed of three parts: retinex loss, denoising loss, and distillation loss; during the teacher model training phase, pairs of low-light images are first processed through retinex loss to extract illumination and radiance, and then through denoising loss to extract denoising prior knowledge. Then, during the student model training phase, distillation loss is used to distill prior knowledge from the teacher model.

Retinex loss: In order to enable accurate processing of the input by an optimal Retinex algorithm, we first remove inappropriate features:

$$\mathscr{L}_P = \|I_1 - i_1\|_2^2, \tag{5}$$

the L_p alters the initial image into a version that is better suited for Retinex decomposition. This transformation is based on the projection image I_1. i_1 is the result of I_1 processed by F-Net. The Retinex loss ensures that the input is optimized for the subsequent Retinex decomposition process.

Based on the paired low-light images and the Retinex theory, we calculate the reflectance consistency loss $\mathscr{L}_E$. Compared to manually crafted priors, $\mathscr{L}_E$ offers greater adaptability and accuracy as it reveals the physical characteristics of objects. In a mathematical representation, $\mathscr{L}_E$ is depicted as:

$$\mathscr{L}_E = \left\| R_{t_1} - R_{t_2} \right\|_2^2, \tag{6}$$

where R_{t_1} and R_{t_2} are the reflectance components of the paired low-light images. $\mathscr{L}_E$ ensures that the network predicts the identical reflectance component for the paired low-light images, as they share the same component.

The decomposition process in Retinex theory typically involves several fundamental constraints that have proven to be highly effective [5]:

$$\begin{aligned}\mathscr{L}_R = &\|R \circ L - i\|_2^2 + \|R - i/stopgrad(L)\|_2^2 \\ &+ \|\nabla L\|_1 + \max(0, L - L_n),\end{aligned} \tag{7}$$

where i represents the projection image, L_n stands for the preliminary estimation of illumination, and ∇ indicates the horizontal and vertical gradients. The reconstruction term $\|R \circ L - i\|_2^2$ ensures that the decomposed elements fulfill the prerequisites for reconstructing the input image. Upon estimating the illumination, the reflectance can be computed by performing a pixel-wise division of the low-light image concerning its illumination map. An additional term $\|R - i/stopgrad(L)\|_2^2$ is added to guide the decomposition. The proposed term $\max(0, L - L_n)$ ensures that when L is less than L_n, the difference is capped at 0, preventing negative values in the loss function. This is used to enforce the constraint $L \leq L_n$. The initial estimation of illumination L_n is calculated by taking the maximum value of the red, green, and blue channels:

$$L_n = \max_{c \in \{R,G,B\}} I^c(x). \tag{8}$$

where $I^c(x)$ represents the values of each pixel x in the red (R), green (G), and blue (B) color channels.

The total Retinex loss function is given by:

$$\mathscr{L}_{Re} = \alpha_1 \mathscr{L}_P + \alpha_2 \mathscr{L}_E + \alpha_3 \mathscr{L}_R, \tag{9}$$

where $\alpha_1, \alpha_2, \alpha_3$ are weights set experimentally to 500, 1, 1, respectively.

Denoising loss: The denoising loss L_D based on the SGN network is designed to propagate the prior knowledge learned from the SGN. It is mathematically represented as:

$$\mathscr{L}_D = \|d - o_t\|_2^2, \tag{10}$$

where d is the prediction from the PD-net and o_t is the one from the D-net. To prevent the prior denoising network from losing essential information in the event of failure, we enforce a similarity between the predictions before and after the D-net using the following loss:

$$\mathscr{L}_O = \|o_m - o_t\|_2^2, \tag{11}$$

where o_m and o_t represent the predictions before and after D-net. The total denoising loss function is:

$$\mathscr{L}_{De} = \beta_1 \mathscr{L}_D + \beta_2 \mathscr{L}_O, \tag{12}$$

where β_1 and β_2 are weights set experimentally to 1 and 1.5, respectively.

Distillation loss: We utilize an offline knowledge distillation approach due to its benefits of flexibility, ease of implementation, and cost-effectiveness compared to other methods. Specifically, we use a similarity-based training method for distillation. This involves measuring the similarity between the Teacher and Student models' outputs and intermediate representations. The loss function encourages the student model to align its outputs and representations with those of the teacher model. To achieve this, the student model is trained to match the teacher model's outputs and intermediate

representations using the following losses based on mean squared error:

$$\mathscr{L}_A = \left\| L_{t_1} - L_s \right\|_2^2 \tag{13}$$

$$\mathscr{L}_B = \left\| R_{t_1} - R_s \right\|_2^2 \tag{14}$$

$$\mathscr{L}_C = \left\| o_t - o_s \right\|_2^2, \tag{15}$$

where $\mathscr{L}_A$ and $\mathscr{L}_B$ calculate the reflectance consistency loss for the illumination and reflectance components decomposed according to the Retinex theory. $\mathscr{L}_C$ represents the prediction similarity between the teacher network and the student network. L_{t1}, R_{t1}, o_t and L_s, R_s, o_s respectively represent the output and intermediate representations of the teacher network and the student network, respectively. The total distillation loss function is given by:

$$\mathscr{L}_{Di} = \eta_1 \mathscr{L}_A + \eta_2 \mathscr{L}_B + \eta_3 \mathscr{L}_C, \tag{16}$$

where the balance coefficients η_1, η_2, and η_3 are all set to 1.

4 Experiments

In this section, we provide a detailed description of our experimental procedures and evaluation processes. Initially, we describe the specific implementation details of experiments, including the assessment datasets employed and the performance metrics utilized. Then, we present a qualitative and quantitative comparative analysis with state-of-the-art methods to evaluate the strengths and characteristics of our proposed approach thoroughly. Finally, We perform detailed ablation studies to assess the effect of each key component on the model's performance, providing a clear understanding of their roles and contributions.

4.1 Implementation Details

We implemented DDR using PyTorch, with training involving random 128×128 image crops and set batch size to 2. We apply Adaptive Moment Estimation [38], initiating the learning rate at 9×10^{-5}, which was halved every 100 epochs over 400 epochs. The adjustment parameter λ was set to 0.2 under normal conditions and 0.14 for darker conditions like those in the LOL dataset. Experiments were conducted on a laptop with an i5-7300HQ CPU, 16 GB RAM, and an NVIDIA GTX 1050Ti GPU, as well as on an NVIDIA Jetson AGX Xavier platform.

4.2 Datasets and Metrics

We extract low-light images from the SICE and LOL datasets for the training of the DDR model. To evaluate our model's performance, we selected a total of 165 images from the SICE and LOL datasets, including 15 images from the official LOL evaluation set, to form our test dataset. Given that both datasets provide ground truth images, we employed a suite of objective metrics, including Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM) [39], Learned Perceptual Image Patch Similarity (LPIPS) [40] and the CIE2000 DeltaE [41], to assess the LIE effectiveness of various methods. Higher PSNR and SSIM scores indicate closer alignment with the reference images, while lower LPIPS and DeltaE values suggest more effective image enhancement. Furthermore, to evaluate the model's capacity for real-time processing, we calculate the computational latency for each image during the testing. Reduced processing durations evidence an optimized real-time processing capability.

4.3 Methods of Comparison

DDR is compared with 13 state-of-the-art LIE methods, which can be categorized into three groups: traditional methods, supervised learning-based methods, and unsupervised methods.

To demonstrate real-time processing capabilities, DDR is also evaluated against these methods in terms of Frames Per Second (FPS). It should be noted that all results are obtained by reproducing the methods using official codes and recommended parameters.

4.4 Quantitative Comparison

Table 1 provides a comprehensive quantitative evaluation of various LIE models across the LOL and SICE datasets, showcasing a comparison of our proposed DDR method with a range of state-of-the-art competitors. The results demonstrates that traditional and unsupervised methods fall short of achieving optimal results, a predictable outcome considering the challenges of developing an effective enhancement algorithm without the guidance of reference imagery. The absence of denoising and prior knowledge of these techniques also limits their adaptability across the varied lighting scenarios typical of real-world settings. Specifically, on the LOL dataset, DDR achieves a PSNR of 19.82 dB, outperforming the second-best unsupervised method by a margin of 0.31 dB and only 0.02 dB less than the top-performing supervised method. For SSIM, DDR attains a score of 0.778, a 0.042 improvement over the runner-up unsupervised method, signifying a substantial enhancement in image quality. The LPIPS value for DDR stands at 0.232, which is the most favorable among unsupervised methods, exceeding the second-best by 0.016. DDR also excels in terms of DeltaE and processing speed, registering a DeltaE of 8.628, an improvement of 2.172 over the second-best unsupervised method and 2.052 over the best-supervised method, indicating the closest match to the ground truth in color accuracy. The DDR algorithm demonstrate superior performance on the SICE dataset. Specifically, it achieves SSIM and DeltaE values of 0.865 and 5.831, respectively, which are 0.022 and 1.897 points higher than the next-best competitor. DDR also sets a new standard for unsupervised methods in terms of PSNR and LPIPS, with scores of 21.51 dB and 0.198, respectively. In terms of processing speed, DDR leads all methods on GPU hardware with a speed of 382.1 fps, outperforming the second-ranked method by 83 fps. Furthermore, on edge devices with limited computational resources, such as AGX, DDR attains the second-highest speed of 199.7 fps. The DDR algorithm's consistent outperformance across various metrics on both the LOL and SICE datasets, as well as its ability to handle real-time processing, underscores its remarkable performance in low-light image enhancement and real-time processing, which highlights its efficiency and efficacy within the domain.

Table 1: Quantitative comparisons with state-of-the-art models on LOL and SICE datasets. "T", "S", and "U" stand for "Traditional Learning", "Supervised Learning", and "Unsupervised Learning" models, respectively. "↑" indicates that higher values are better, while "↓" indicates that lower values are preferable. **Red**, **blue**, and **green** indicate the first, second and third place

Method	Tpye	LOL				SICE				Speed (fps)	
		PSNR↑	SSIM↑	LPIPS↓	DeltaE↓	PSNR↑	SSIM↑	LPIPS↓	DeltaE↓	GPU	AGX
STAR [3] TIP'20	T	12.91	0.518	0.366	23.46	15.17	0.727	0.246	16.35	1.356	0.874
SDD [7] TMM'20	T	13.34	0.637	0.743	21.83	15.35	0.741	0.232	16.08	0.339	0.212
MBLLEN [17] BMVC'18	S	17.86	0.727	**0.225**	13.68	13.64	0.632	0.297	18.60	0.207	–
RetinexNet [2] BMVC'18	S	17.55	0.648	0.379	12.69	19.89	0.783	0.276	8.715	–	2.228
GLADNet [42] FG'18	S	**19.72**	0.680	0.321	12.28	18.98	0.837	0.203	8.947	0.324	–
KinD [43] MM'19	S	17.65	**0.775**	**0.171**	12.49	21.10	0.838	**0.195**	8.009	0.405	–
URetinexNet [24] CVPR'22	S	**19.84**	**0.826**	**0.128**	**10.65**	**21.64**	**0.843**	**0.192**	**7.728**	2.760	1.743
ZeroDCE [8] CVPR'20	U	14.86	0.559	0.335	18.81	18.69	0.810	0.207	11.93	10.79	9.633

(Continued)

Table 1 (continued)

Method	Tpye	LOL				SICE				Speed (fps)	
		PSNR↑	SSIM↑	LPIPS↓	DeltaE↓	PSNR↑	SSIM↑	LPIPS↓	DeltaE↓	GPU	AGX
RRDNet [44] ICME'20	U	11.40	0.457	0.362	26.43	13.28	0.678	0.221	19.64	–	–
RUAS [27] CVPR'21	U	18.23	0.717	0.270	16.85	13.18	0.734	0.363	16.81	**132.0**	27.64
EnlightenGAN [10] TIP'21	U	17.48	0.651	0.322	15.50	18.73	0.822	0.216	10.42	11.34	2.238
PairLIE [5] CVPR'23	U	19.51	0.736	0.248	**10.80**	**21.32**	**0.840**	0.216	**7.835**	**299.1**	**67.87**
LOLEVT [9] CVPR'24	U	16.44	0.701	0.350	–	–	–	–	–	–	**199.7**
DDR (Ours)	U	**19.82**	**0.778**	0.232	**8.628**	**21.51**	**0.865**	**0.198**	**5.831**	**382.1**	**143.4**

4.5 Qualitative Comparison

Fig. 2 presents a qualitative comparison with state-of-the-art LIE methods. Our findings can be summarized as follows:

1) The method we introduced delivers pleasing visual outcomes regarding luminosity, hue, tonal balance, and overall natural appearance. In contrast, alternative methods encounter difficulties when dealing with severely dark lighting scenarios. 2) Although it can be seen from Table 1 that supervised learning methods like KinD, GLADNet, and URetinexNet perform well on the LOL and SICE datasets, supervised learning models may encounter limitations in generalization ability because of their high sensitivity to data distribution.

Figure 2: Visual comparison with state-of-the-art LIE methods on MEF unsupervised dataset. Zoom in to obtain the optimal view. SDD [7], GLADNet [42], KinD [43], URetinext [24], Zero-DCE [8], RUAS [27], SCI [45], EnlightenGAN [10] and Ours. The visual results highlight DDR's superiority in enhancing luminosity, color balance, and texture details, particularly in severely low-light areas

In Fig. 3, we further demonstrate noise suppression examples. Combined with Table 1, it can be observed that DDR, while introducing manual prior knowledge about noise, actually improves its real-time processing performance. In this case, our method successfully suppresses sensor noise in the dark areas, resulting in sharp and authentic images. Conversely, rival approaches often either increase noise or are unable to accurately correct color and contrast, resulting in subpar image quality, especially when considering real-time processing performance.

Figure 3: The visual comparisons of the noise reduction. For the best view, please zoom in. RUAS [27], SCI [45], EnlightenGAN [10], Zero-DCE [8], PairLIE [5] and Ours. Our result is visually satisfying with no obvious noise

4.6 Decomposition Visualization

To evaluate the performance of our model, we conduct a visual demonstration of the reflectance and illumination components. As illustrated in Fig. 4, the reflectance component contains a wealth of texture and detail, whereas the illumination component appears to be segmented and lacks textural features, indicating that the DDR method is capable of effectively separating the components of low-light images. We use different correction factors to demonstrate the enhancement results. As λ increases, the brightness gradually reduces. Specifically, when λ exceeds 0.4, the enhanced images exhibit noticeable under-enhancement effects, whereas when λ falls below 0.2, over-enhancement effects occur. In this work, the default value of λ is set to 0.2. Please note that adjustments to λ can be made based on user preferences during the testing phase.

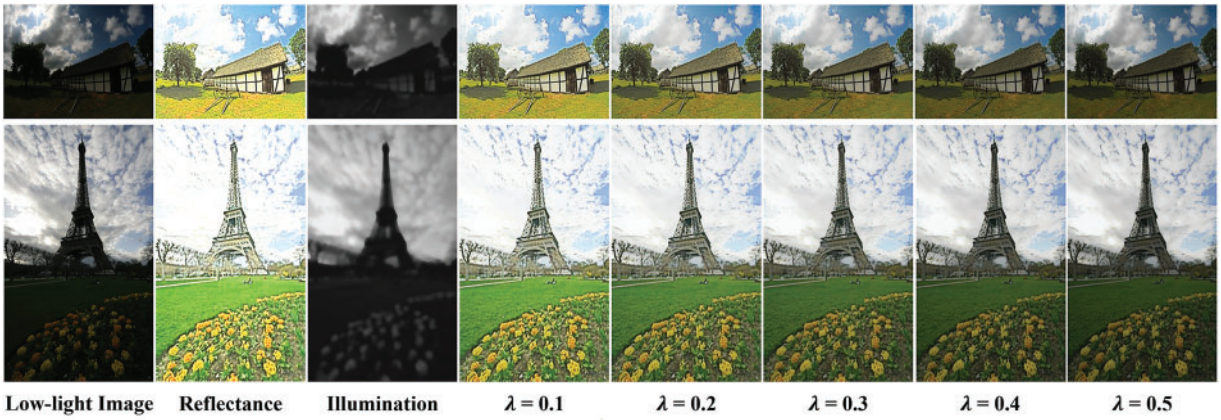

Figure 4: The graphical representation of Retinex decomposition is showcased, with the improved outcomes displayed across a range of correction coefficients. The standard value for λ is designated as 0.2. Users have the flexibility to modify the λ parameter to suit their specific requirements

4.7 Ablation Studies

Ablation studies are conducted under various configurations to understand the impact of individual elements on outcomes. The subsequent modifications are made to the initial DDR: Configuration A: $\mathscr{L}_{De}$ is removed. Configuration B: $\mathscr{L}_{Di}$ is removed. Configuration C: Prior terms are removed, i.e., both $\mathscr{L}_{De}$ and $\mathscr{L}_{Di}$ are eliminated.

Table 2 presents the results of the ablation study conducted on the LOL and SICE datasets. The data indicate that our proposed method significantly outperforms Configurations A and C in terms of LIE on both datasets, thereby validating the superiority of incorporating a denoising subnetwork to learn adaptive priors. This suggests that the denoising subnetwork effectively learns adaptive priors from low-light images, thereby improving the quality of image enhancement. It is noteworthy that on the LOL dataset, our method is slightly inferior to Configuration B, with a minimal difference of 0.01 dB in PSNR, 0.003 in SSIM, and 0.004 in LPIPS. In contrast, on the SICE dataset, our method shows a particular improvement over Configuration B, with increases of 0.08 dB in PSNR, 0.014 in SSIM, and 0.007 in LPIPS, which may be attributed to the enhanced generalization capability of the model due to knowledge distillation. This improvement may be attributed to the distillation process optimizing the model structure, reducing the number of parameters and computational complexity, thus speeding up inference. Regarding real-time processing performance, the DDR method excels. Specifically, on the LOL dataset, the fps is 139.1, which is 74.7 fps higher than the optimal Configuration B, corresponding to a 116.0% improvement. On the SICE dataset, the DDR method achieves an even more impressive processing speed of 147.7 fps, which is 79.9 fps quicker than Configuration B, translating to a 117.8% increase in speed. These results indicate that knowledge distillation significantly enhances the real-time processing performance of DDR.

Table 2: The quantitative outcomes of ablation experiments on the LOL and SICE datasets are presented. The top results are highlighted in bold

Dataset	Configuration	Loss functions		PSNR↑	SSIM↑	LPIPS↓	Speed (fps)
		$\mathscr{L}_{Di}$	$\mathscr{L}_{De}$				AGX
LOL	A	✓	×	19.31	0.766	0.323	132.6
	B	×	✓	**19.85**	**0.781**	**0.228**	52.45
	C	×	×	19.31	0.736	0.257	64.40
	DDR	✓	✓	19.84	0.778	0.232	**139.1**
SICE	A	✓	×	21.23	0.856	0.250	141.9
	B	×	✓	21.43	0.851	0.205	71.34
	C	×	×	21.32	0.840	0.216	67.87
	DDR	✓	✓	**21.51**	**0.865**	**0.198**	**147.7**

Fig. 5a demonstrates the impact of knowledge distillation on the processing time for each image when deployed on AGX devices. The blue line (Ours) represents the processing time with distillation, while the orange line (Baseline) indicates the time without distillation (i.e., Configuration B). The figure clearly shows that the processing time with distillation is markedly shorter than the baseline method. Fig. 5b demonstrates the visual comparison of LIE performance on the LOL dataset under different ablation experiment Configurations. It can be observed that the visual effect of DDR is significantly better than Configurations A and C and slightly lower than Configuration B, which is consistent with the experimental parameters in Table 2. In summary, this distilled representation not only maintains the performance of image enhancement but also significantly reduces the computational time required for inference. The results of Configuration C further demonstrate that the removal of the prior term leads to a substantial degradation in model performance.

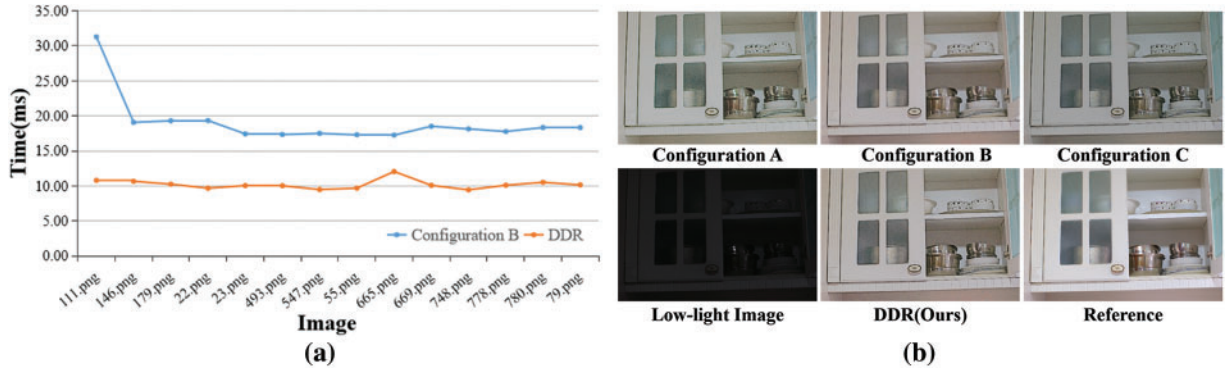

Figure 5: Comparison of ablation experiments. (a) Real-time performance on the LOL dataset using AGX platform, Baseline refers to the Configuration B; (b) Visual comparison of ablation experiments

4.8 Discussions

We evaluated the impact of DDR on image quality across three different color spaces: YCbCr, HSV, and CIE-Lab. Table 3 presents the results, which indicate that images in the YCbCr and CIE-Lab color spaces showed significant improvements in both PSNR and SSIM after enhancement, increasing from 11.8074 dB and 0.6401 to 17.5135 dB and 0.7799, respectively. This suggests that the enhancement techniques effectively improved the visual quality and structural similarity of images in these color spaces. However, the enhancement effects in the HSV color space were not satisfactory, with a decrease in PSNR and SSIM, from 11.8074 to 10.7047 dB and from 0.6401 to 0.4157, respectively. This may be attributed to the sensitivity of the HSV color space to changes in brightness. The LPIPS increased slightly in the YCbCr and CIE-Lab color spaces but significantly in the HSV color space, indicating a a more considerable perceptual difference between the enhanced and original images. These findings highlight the critical role of color space selection in the performance of image enhancement techniques. Future research can further explore the adaptability and optimization methods of different color spaces to enhance the performance of low-light image enhancement technologies.

Table 3: Image enhancement results in different color spaces

Color space	Original			Enhanced		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
YCbCr	11.8074	0.6401	0.6617	17.5135	0.7799	0.7842
HSV	11.8074	0.6401	0.6617	10.7047	0.4157	1.0266
CIE-Lab	11.7255	0.6479	0.6674	17.5135	0.7799	0.7842

The limitations of this approach include constrained noise handling effectiveness in complex scenes and challenges in achieving real-time performance under highly resource-limited conditions. Future research directions could focus on dynamic noise adaptation, hybrid supervised-unsupervised denoising techniques, and optimizing the model through pruning or quantization for deployment on edge devices. The proposed method's potential economic impact lies in its reduced reliance on costly labeled data, making it suitable for low-light enhancement in sectors such as surveillance, automotive, and healthcare. It also supports cost-effective deployment on low-power devices, enhancing the usability of image data.

5 Conclusions

In this work, we propose an unsupervised LIE method to address the challenges of noise and detail loss in poorly lit environments by incorporating prior knowledge from a pre-trained denoising model into the Retinex-based framework. Moreover, knowledge distillation is employed to refine the model. Experimental results demonstrate that our method achieves a PSNR of 19.82 dB and an SSIM of 0.778 on the LOL dataset, indicating a significant improvement in image quality. Furthermore, our model achieves a processing speed of 382.1 fps, outperforming existing unsupervised methods. These findings demonstrate that our solution offers a practical and efficient approach to LIE tasks, setting a benchmark in both performance and applicability. However, there is still room for improvement when enhancing low-light images in other color spaces or under conditions of severe resource constraints. Future research could focus on exploring the adaptability and optimization methods of different color spaces and optimizing the model through pruning or quantization for deployment on edge devices.

Acknowledgement: We express our deep appreciation to our families and friends for their constant support and encouragement.

Funding Statement: Thanks to the support by the Guangxi Natural Science Foundation (Grant No. 2024GXNSFAA010484), and the National Natural Science Foundation of China (No. 62466013), this work has been made possible.

Author Contributions: Study conception and design: Wenkai Zhang, Xingzhe Wang, Shuiwang Li; data collection: Wenkai Zhang, Xianming Liu; analysis and interpretation of results: Wenkai Zhang, Hao Zhang; draft manuscript preparation: Wenkai Zhang, Hao Zhang, Xiaoyu Guo. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: Data openly available in a public repository. Code is available at: https://github.com/qw631399/DDR (accessed on 13 November 2024). The other used materials will be made available on request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

[1] L. H. Pham, D. N. -N. Tran, and J. W. Jeon, "Low-light image enhancement for autonomous driving systems using DriveretiNex-Net," in *2020 IEEE Int. Conf. Consum. Electr.-Asia (ICCE-Asia)*, Seoul, Republic of Korea, 2020, pp. 1–5. doi: 10.1109/ICCE-Asia49877.2020.9277442.

[2] C. Wei *et al.*, "Deep retinex decomposition for low-light enhancement," 2018, *arXiv:1808.04560*.

[3] J. Xu, Y. K. Hou, D. W. Ren, L. Liu, F. Zhu and M. Y. Yu, "STAR: A structure and texture aware retinex model," *IEEE Trans. Image Process.*, vol. 29, pp. 5022–5037, 2020. doi: 10.1109/TIP.2020.2974060.

[4] J. Cai, S. Gu, and L. Zhang, "Learning a deep single image contrast enhancer from multi-exposure images," *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 2049–2062, 2018. doi: 10.1109/TIP.2018.2794218.

[5] Z. Q. Fu, Y. Yang, X. T. Tu, Y. Huang, X. H. Ding and K. K. Ma, "Learning a simple low-light image enhancer from paired low-light instances," in *Proc. 2023 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Vancouver, BC, Canada, 2023, pp. 22252–22261.

[6] S. Gu, Y. W. Li, L. Van Gool, and R. Timofte, "Self-guided network for fast image denoising," in *Proc. 2019 IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Seoul, Republic of Korea, 2019, pp. 2511–2520.

[7] S. Hao, X. Han, Y. Guo, X. Xu, and M. Wang, "Low-light image enhancement with semi-decoupled decomposition," *IEEE Trans. Multimed.*, vol. 22, no. 12, pp. 3025–3038, 2020. doi: 10.1109/TMM.2020.2969790.

[8] C. Guo, C. Li, J. Guo, C. Change Loy, J. Hou and S. Kwong, "Zero-reference deep curve estimation for low-light image enhancement," in *Proc. 2020 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, USA, 2020, pp. 1780–1789.

[9] S. M. A. Sharif, A. Myrzabekov, N. Khujaev, R. Tsoy, S. Kim and J. Lee, "Learning optimized low-light image enhancement for edge vision tasks," in *Proc. 2024 IEEE/CVF Conference on Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Seattle, WA, USA, 2024, pp. 6373–6383.

[10] Y. F. Jiang, X. Y. Gong, D. Liu, Y. Cheng, C. Fang and X. H. Shen, "EnlightenGAN: Deep light enhancement without paired supervision," *IEEE Trans. Image Process.*, vol. 30, pp. 2340–2349, 2021. doi: 10.1109/TIP.2021.3051462.

[11] W. Lee, J. Lee, D. Kim, and B. Ham, "Learning with privileged information for efficient image super-resolution," in *Comput. Vis.-ECCV 2020: 16th European Conf.*, Glasgow, UK, Springer, 2020, pp. 465–482.

[12] Z. Wang, D. Zhao, and Y. Cao, "Image quality enhancement with applications to unmanned aerial vehicle obstacle detection," *Aerospace*, vol. 9, no. 12, 2022, Art. no. 829. doi: 10.3390/aerospace9120829.

[13] Y. Zhu, L. Wang, J. Yuan, and Y. Guo, "Diffusion model based low-light image enhancement for space satellite," 2023, *arXiv:2306.14227*.
[14] J. Guo, J. Ma, F. García-Fernández Ángel, Y. Zhang, and H. Liang, "A survey on image enhancement for low-light images," *Heliyon*, vol. 9, no. 4, 2023, Art. no. e14558. doi: 10.1016/j.heliyon.2023.e14558.
[15] W. Song, M. Suganuma, X. Liu, N. Shimobayashi, D. Maruta and T. Okatani, "Matching in the dark: A dataset for matching image pairs of low-light scenes," in *Proc. IEEE Int. Conf. Comput. Vis.*, Montreal, QC, Canada, 2021, pp. 6029–6038.
[16] K. G. Lore, A. Akintayo, and S. Sarkar, "LLNet: A deep autoencoder approach to natural low-light image enhancement," *Pattern Recognit.*, vol. 61, pp. 650–662, 2017. doi: 10.1016/j.patcog.2016.06.008.
[17] F. Lv, F. Lu, J. Wu, and C. Lim, "MBLLEN: Low-light image/video enhancement using CNNs," in *Brit. Mach. Vis. Conf.*, Northumbria University, 2018.
[18] C. Li, J. Guo, F. Porikli, and Y. Pang, "LightenNet: A convolutional neural network for weakly illuminated image enhancement," *Pattern Recognit. Lett.*, vol. 104, pp. 15–22, 2018. doi: 10.1016/j.patrec.2018.01.010.
[19] C. Chen, Q. Chen, J. Xu, and V. Koltun, "Learning to see in the dark," in *Proc. 2018 IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA, 2018, pp. 3291–3300.
[20] G. -H. Park, H. -H. Cho, and M. -R. Choi, "A contrast enhancement method using dynamic range separate histogram equalization," *IEEE Trans. Consum. Electron.*, vol. 54, no. 4, pp. 1981–1987, 2008. doi: 10.1109/TCE.2008.4711262.
[21] C. Lee, C. Lee, and C. -S. Kim, "Contrast enhancement based on layered difference representation of 2D histograms," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 5372–5384, 2013. doi: 10.1109/TIP.2013.2284059.
[22] X. Guo, Y. Li, and H. Ling, "LIME: Low-light image enhancement via illumination map estimation," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 982–993, 2016. doi: 10.1109/TIP.2016.2639450.
[23] M. Li, J. Liu, W. Yang, X. Sun, and Z. Guo, "Structure-revealing low-light image enhancement via robust retinex model," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2828–2841, 2018. doi: 10.1109/TIP.2018.2810539.
[24] W. Wu, J. Weng, P. Zhang, X. Wang, W. Yang and J. Jiang, "Uretinex-Net: Retinex-based deep unfolding network for low-light image enhancement," in *Proc. 2022 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, New Orleans, LA, USA, 2022, pp. 5901–5910.
[25] X. Xu, R. Wang, C. -W. Fu, and J. Jia, "SNR-aware low-light image enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, New Orleans, LA, USA, 2022, pp. 17714–17724.
[26] Z. Zhang, H. Zheng, R. Hong, M. Xu, S. Yan and M. Wang, "Deep color consistent network for low-light image enhancement," in *Proc. 2022 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, New Orleans, LA, USA, 2022, pp. 1899–1908.
[27] R. Liu, L. Ma, J. Zhang, X. Fan, and Z. Luo, "Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement," in *Proc. 2021 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Nashville, TN, USA, 2021, pp. 10561–10570.
[28] S. Yang, M. Ding, Y. Wu, Z. Li, and J. Zhang, "Implicit neural representation for cooperative low-light image enhancement," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2023, pp. 12918–12927.
[29] K. A. Hashmi, G. Kallempudi, D. Stricker, and M. Z. Afzal, "FeatEnHancer: Enhancing hierarchical features for object detection and beyond under low-light vision," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2023, pp. 6725–6735.
[30] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," 2015, *arXiv:1503.02531*.
[31] T. Chen, I. Goodfellow, and J. Shlens, "Net2Net: Accelerating learning via knowledge transfer," 2015, *arXiv:1511.05641*.
[32] L. Zhang, J. Song, A. Gao, J. Chen, C. Bao and K. Ma, "Be your own teacher: Improve the performance of convolutional neural networks via self distillation," in *Proc. 2019 IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Seoul, Republic of Korea, 2019, pp. 3713–3722.
[33] S. Zagoruyko and N. Komodakis, "Paying more attention to attention: Improving the performance of convolutional neural networks via attention transfer," 2016, *arXiv:1612.03928*.

[34] Y. Liu, K. Chen, C. Liu, Z. Qin, Z. Luo and J. Wang, "Structured knowledge distillation for semantic segmentation," in *Proc. 2019 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Long Beach, CA, USA, 2019, pp. 2604–2613.
[35] Q. Gao, Y. Zhao, G. Li, and T. Tong, "Image super-resolution using knowledge distillation," in *Asian Conf. Comput. Vis.*, Springer, 2018, pp. 527–541.
[36] Z. He, T. Dai, J. Lu, Y. Jiang, and S. -T. Xia, "Fakd: Feature-affinity based knowledge distillation for efficient image super-resolution," in *2020 IEEE Int. Conf. Image Process. (ICIP)*, Abu Dhabi, United Arab Emirates, 2020, pp. 518–522. doi: 10.1109/ICIP40778.2020.9190917.
[37] Z. Li, Y. Wang, and J. Zhang, "Low-light image enhancement with knowledge distillation," *Neurocomputing*, vol. 518, pp. 332–343, 2023. doi: 10.1016/j.neucom.2022.10.083.
[38] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[39] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004. doi: 10.1109/TIP.2003.819861.
[40] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. 2018 IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 586–595.
[41] G. Sharma, W. Wu, and E. N. Dalal, "The ciede2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations," *Color Res.*, vol. 30, no. 1, pp. 21–30, 2005. doi: 10.1002/(ISSN)1520-6378.
[42] W. Wang, C. Wei, W. Yang, and J. Liu, "GLADNet: Low-light enhancement network with global awareness," in *2018 13th IEEE Int. Conf. Automatic Face Gesture Recognit. (FG 2018)*, Xi'an, China, 2018, pp. 751–755. doi: 10.1109/FG.2018.00118.
[43] Y. Zhang, J. Zhang, and X. Guo, "Kindling the darkness: A practical low-light image enhancer," in *Proc. 27th ACM Multimed.*, 2019, pp. 1632–1640.
[44] A. Zhu, L. Zhang, Y. Shen, Y. Ma, S. Zhao and Y. Zhou, "Zero-shot restoration of underexposed images via robust retinex decomposition," in *2020 IEEE Int. Conf. Multimed. Expo (ICME)*, IEEE, 2020, pp. 1–6.
[45] L. Ma, T. Ma, R. Liu, X. Fan, and Z. Luo, "Toward fast, flexible, and robust low-light image enhancement," in *Proc. 2022 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, New Orleans, LA, USA, 2022, pp. 5637–5646.