Low Light Image Enhancement

A

Project Report

submitted

in partial fulfillment

for the award of the Degree of

Bachelor of Technology

in Department of Computer Science

Supervisor:

Dr. Satyanarayan Tazi

Submitted By:

Gaurav Rathore (20EEACS036)

Gaurav Singh (20EEACS037)

Gautam Giri Goswami (20EEACS038)

Shubhangi Tyagi (20EEACS097)

Department of Computer Science & Engineering

Engineering College Ajmer

June, 2024

Candidate's Declaration

I hereby declare that the work, which is being presented in the Project, entitled **"Low Light Image Enhancement"** in partial fulfillment for the award of Degree of "*Bachelor* of Technology" in Department of **Computer Science,** Engineering College Ajmer, Bikaner Technical University is a record of my own investigations carried under the Guidance of **Dr. S.N. Tazi**, Department of Computer Science & Engineering, Engineering College Ajmer. I have not submitted the matter presented in this report anywhere for the award of any other Degree.

(Gaurav Rathore)

University Roll Number: 20EEACS036

(Gaurav Singh)

University Roll Number: 20EEACS037

(Gautam Giri Goswami)

University Roll Number: 20EEACS038

(Shubhangi Tyagi)

University Roll Number: 20EEACS097

ENGINEERING COLLEGE AJMER
(AN AUTONOMOUS INSTITUTION OF GOVT. OF RAJASTHAN)
NH-8, Barliya Circle, Ajmer (Rajasthan) -305025
Department of Computer Science & Engineering

CERTIFICATE

This is to certify that ***Gaurav Rathore, Gaurav Singh, Gautam Giri Goswami, Shubhangi Tyagi*** of VIII Semester, B.Tech (***Computer Science***) 2023-24, has submitted the Project titled **"*Low Light Image Enhancement*"** in partial fulfilment for the award of the degree of Bachelor of Technology under Bikaner Technical University, Bikaner.

Date: **Dr. Satyanarayan Tazi**

(Supervisor)

ACKNOWLEDGMENET

We take this opportunity to express our gratitude to all those people who have been directly and indirectly with us during the completion of this Project. We pay thanks to **Dr. S.N. Tazi** who has given guidance and a light to us during this major project. His versatile knowledge about **“Low Light Image Enhancement”** has eased us in the critical times during the span of this Project.

We acknowledge here out debt to those who contributed significantly to one or more steps. I take full responsibility for any remaining sins of omission and commission.

Gaurav Rathore

Gaurav Singh

Gautam Giri Goswami

Shubhangi Tyagi

(Computer Science)

ABSTRACT

Low-light image enhancement (LIE) is essential in improving the visibility and quality of images captured under poor lighting conditions, which is crucial for various computer vision applications. Existing methods, including traditional histogram-based and Retinex-based approaches, often depend on handcrafted priors, which may not be adaptive enough for diverse scenes and lighting conditions. On the other hand, supervised learning techniques require extensive labelled datasets that are challenging to acquire. This report introduces PairLIE, an innovative reference-free LIE approach that leverages paired low-light images to learn adaptive priors in a data-driven manner, minimizing the need for handcrafted features and reference images.

The proposed method incorporates a simple yet effective network architecture comprising L-Net, R-Net, and P-Net to estimate illumination and reflectance components based on the Retinex theory. PairLIE employs a combination of projection loss, reflectance consistency loss, and Retinex loss to ensure accurate decomposition and enhancement of low-light images. The methodology involves collecting low-light image pairs from datasets such as SICE and LOL, and training the network using these pairs to improve performance and generalization. Extensive experiments conducted on public benchmarks reveal that PairLIE significantly outperforms state-of-the-art unsupervised methods and achieves competitive results compared to supervised approaches. The experimental results highlight PairLIE's effectiveness in enhancing image brightness, color consistency, and overall visual quality without relying on extensive handcrafted priors or reference images. This research underscores the potential of utilizing paired low-light images to develop robust and adaptive low-light image enhancement solutions, paving the way for future advancements in this field.

CONTENT

INTRODUCTION

Images captured in low-light environments often suffer from multiple distortions such as low contrast, poor visibility, and sensor noise. These issues make low-light images unsatisfactory for both human visualization and subsequent computer vision tasks. The challenges associated with low-light image enhancement (LIE) have led to significant research efforts aimed at correcting contrast, uncovering textures, and removing sensor noise.

Traditional LIE methods can be broadly categorized into histogram-based and Retinex-based approaches. Histogram-based techniques enhance image contrast by redistributing the luminous intensity across the histogram. Retinex-based methods, on the other hand, decompose an observed image into illumination and reflectance components. The reflectance component, which represents the physical properties of objects, remains consistent under different lighting conditions, making Retinex-based methods particularly appealing due to their alignment with human color perception.

Despite the success of these traditional methods, they often fall short in complex natural scenes. Recent advancements in learning-based LIE algorithms have shown great promise, particularly those leveraging paired low-light and normal-light images. However, acquiring high-quality normal-light reference images is often time-consuming and expensive, limiting the practicality of these methods. To address this, unsupervised and zero-shot LIE approaches have been developed, which either train networks using low-light samples or optimize using the test image itself. These approaches, however, still rely heavily on handcrafted priors and often struggle with high-quality enhancement due to the limited information available in single low-light images.

Our project overcome these limitations which leverages paired low-light images for training. Unlike traditional methods that depend on normal-light reference images, our approach uses pairs of low-light images to provide richer information for the network. This not only simplifies the data collection process but also enhances the adaptability and robustness of the network by reducing the dependency on handcrafted priors.

Our method employs the Retinex theory to decompose low-light images into illumination and reflectance components. By projecting the original images to remove inappropriate features before applying the Retinex decomposition, we can avoid sub-optimal estimations and achieve more accurate results. This method represents a practical and efficient solution for low-light image enhancement.

1.1 Overview

Low-light image enhancement (LIE) is a crucial area of research within computer vision, aiming to improve the visibility and quality of images captured in suboptimal lighting conditions. These enhancements have practical applications in various fields, such as photography, surveillance, medical imaging, and autonomous driving. Traditional LIE methods often rely on handcrafted priors and supervised learning techniques, which can be limited by their dependency on reference images and predefined features. In our project we have use the approach that leverages paired low-light images to learn adaptive priors, reducing the reliance on handcrafted features and enhancing generalization across diverse scenes.

1.2 Motivation

The motivation for this project stems from the limitations of existing LIE methods. Traditional approaches often struggle with varying lighting conditions and scene complexities due to their reliance on handcrafted priors. Supervised methods, while effective, require extensive labeled datasets, which are costly and time-consuming to obtain. To address these challenges, we explore a data-driven, unsupervised approach that can learn from the inherent relationships within low-light image pairs. This method aims to improve LIE performance by automatically adapting to different scenes and lighting conditions, thus providing a more robust and scalable solution.

1.3 Objective

The primary objective of this project is to develop a low-light image enhancement method that:

1. Utilizes paired low-light images to learn adaptive priors for the Retinex decomposition.
2. Reduces dependence on handcrafted features and reference images.

3. Achieves competitive performance with state-of-the-art supervised and unsupervised LIE methods.
4. Enhances generalization across various scenes and lighting conditions through a simpler and more efficient network design.

CHAPTER 2
LITERATURE SURVEY

Low-light image enhancement (LIE) is a crucial task in computer vision, aimed at improving the visibility and quality of images captured in suboptimal lighting conditions. Over the years, extensive research has been devoted to developing methods that enhance the brightness, contrast, and overall quality of low-light images while preserving details and minimizing artifacts.

Conventional machine learning (CML) is a common machine learning approach that uses programmed rules and models to identify patterns and make decisions. CML algorithms use statistical models and traditional programming techniques to process data and make predictions[1]

2.1 Low-Light Image Enhancement

Low-light image enhancement refers to the process of improving the quality and visibility of images captured in low-light conditions. When a camera captures an image in low-light situations, several challenges arise, including reduced brightness, increased noise, and loss of details and colors.[2] Low-light image enhancement aims to address these challenges and produce visually pleasing and informative images.

2.1.1 Traditional and Plain Methods

Plain methods usually enhanced image by histogram. Traditional Methods commonly depend on the application of the Retinex theory, which decom posed the lights into illumination and reflections. For example, refine the initial illumination map to optimize lighting details by imposing a structure prior.[2] Regrettably, existing methodologies are inadequate in effectively eliminating noise artifacts and producing accurate color mappings, rendering incapable of achieving the desired level of precision and finesse in LLIE task.

2.1.2 Deep Learning Methods

Deep learning-based approaches have been widely used in LLIE task. For instance, Retinex enhancing images based on Retinex theory. [9] RUAS unrolled with architecture search to construct lightweight yet effective LLIE network. SNR-Aware present a collectively exploiting Signal-to-Noise-Ratio-aware trans formers to dynamically enhance pixels with spatial-varying operations. Still, all of these methods are recovered on sRGB space, which is not only inaccurately controlled in terms of brightness, but also biased in terms of color.

2.2 Color Space

A color space is a specific organization of colors that allows us to represent and manipulate color in a visual medium, such as digital images or printed materials. In simpler terms, it's a mathematical model that describes how colors can be represented as numerical values.[11]

Color spaces define a range of colors that can be represented and provide a method for interpreting those colors. Different color spaces have different properties and are used for different purposes, depending on factors such as the medium of display, the capabilities of the imaging device, and the requirements of the application.

2.2.1 RGB Color Space

Any additive color space based on RGB color model belongs to the RGB color space. Currently the most common used is the standard-RGB (sRGB) color space.[11]

For the same principle as visual recognition by the human eye, sRGB is widely used in digital imaging devices. Nevertheless, sRGB is coupled in three axis and not suitable for enhancement as presented.

2.2.2 HSV and HSL Color Space

Hue, Saturation and Value or Lightness color space is a method of representing points in an RGB color model in a cylindrical coordinate system. Indeed, it does de coupled brightness and color of the image from RGB channels. However, the inherent hue axis color discontinuity and non-mono-mapped pure black planes pose significant challenges when attempting to enhance the image in HSV color space, resulting in the emergence of highly pronounced artifacts.

2.3 HVI Color Space

To sort out the aforementioned color space challenges, we first present a trainable Horizontal/Vertical-Intensity (HVI) color space.[13] It consists of three trainable parameters and a custom training function that can adapt to the photosensitive characteristics and color sensitivities of the dataset. Specifically, our focus lies on developing a mono-mapping trans form that enables the conversion between sRGB and HVI.

2.3.1 Intensity Map

Any single sRGB image $I \in RH\times W\times 3$ can be decomposed into three image $Ic \in RH\times W$ where $c \in \{R,G,B\}$. There we denote the pixel point light intensity by $Imax = max$ $c\in\{R,G,B\}$ (Ic) which represent intensity map.

2.3.2 HV Color Map

We model a HV Color map as a plane to quantify color reflectance map,[10] which can be trained. Inspired by the weakness of the tiny value of low-lights in sRGB

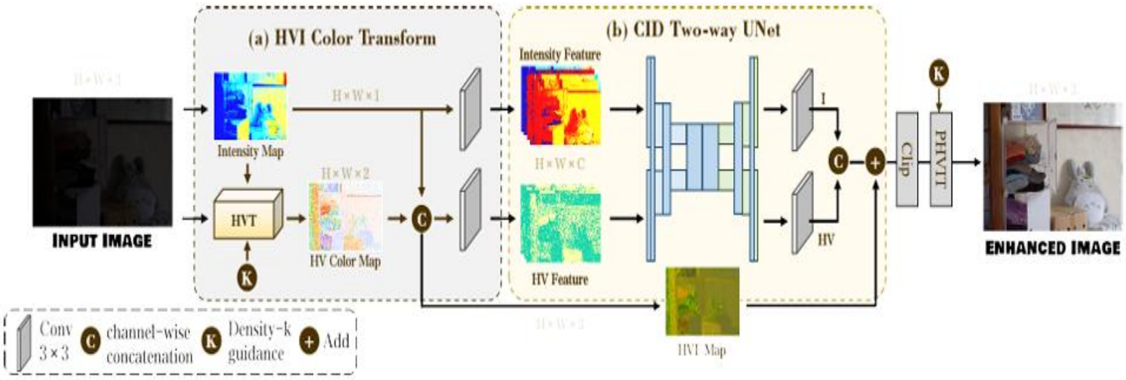


Figure 2.1 – The overview CID Net pipeline of our method.

- In figure 2.1 we see, HVI Color Transform (HVIT) :It takes an sRGB image as input and generates HV feature , Intensity feature , and HVI map as outputs. CID Two-way UNet : This module performs the main processing , utilizing a Two-way UNet architecture . Lastly , we apply Perceptual – inverse HVI Transform (PHVIT) , which takes a light – up HVI map as input and transforms it to an sRGB enhanced image.[3]

2.3.3 Trainable HVI Map

We concatenate $\hat{H}$, $\hat{V}$, $Imax$ in Eq. (1,7) to novel $IHVI \in RH\times W\times 3$map. Simultaneously , HVI map enables a one-to one correspondence between each color point in the sRGB color space and the HVI color space , with the added advantage of ensuring reversibility of the transformation . This map can be recognized by computer perceptually . Not only solve it the problem of multi – mapping of pure black planes and inter mitten color axis , but moreover , it can adapt itself to a variety of tasks by way of machine learning for more functionality. We ponder there are still many potential application possibilities to explore in this color space. For instance , HVI with $T(x)=-4x(x-1)$ is a color space which filter the red related colors . Yet in this paper , we set $T(x)=1, \gamma G= 1\ 6, \gamma B=1\ 3$ to simplify the process and to satisfy image enhance assignment

Algorithm Used

In this, we discuss the Color and Intensity Decoupling Network (CIDNet) and its use of the Horizontal/Vertical-Intensity (HVI) color space for low-light image enhancement. The CIDNet incorporates a Lightweight Cross-Attention (LCA) module to facilitate interaction between image structures and content information in both brightness and color branches, while suppressing noise in low-light images. The proposed HVI color space is designed to address the instability and sensitivity issues in existing color spaces when used for low-light image enhancement. It also presents experimental results demonstrating the superior performance of CIDNet in enhancing low-light images while preserving natural colors.

The algorithm used in this Model is the Color and Intensity Decoupling Network (CIDNet) for low-light image enhancement, which is based on the Horizontal/Vertical-Intensity (HVI) color space and incorporates the Lightweight Cross-Attention (LCA) module.

2.4 CID Net

CIDNet, or Color and Intensity Decoupling Network, is a novel method for low-light image enhancement that is based on the Horizontal/Vertical-Intensity (HVI) color space. It consists of two pathways dedicated to processing the decoupled image brightness and color in the HVI space. The CIDNet architecture incorporates a Two-way UNet with six Lightweight Cross-Attention (LCA) modules, which facilitate interaction between image structure and content information in both brightness and color branches, while also suppressing noise in low-light images.

The HVI color space, upon which CIDNet is built, effectively decouples image brightness and color, adapts to various image illumination scales, and incorporates trainable representation parameters and a trainable function. The CIDNet's Two-way network concurrently processes the brightness and color of low-light images, and the bidirectional LCA module enhances the interaction between the structures of images contained in the brightness and color branches.

CID Net has been extensively evaluated through 5 quantitative and qualitative experiments, demonstrating its superior performance over state-of-the-art methods across 2 datasets. It effectively enhances the brightness of low-light images while preserving their natural colors and exhibits relatively small parameter and computational loads.

In summary, CID Net is a powerful and efficient network for low-light image enhancement, leveraging the innovative HVI color space and LCA module to achieve superior results in enhancing low-light images.

Our goal at Color and Intensity Decoupling Network (CID Net) is to design a color feature decoupling Transformer that is lightweight and follows the HVI color space for LLIE task.

2.4.1 Pipeline

Allow-lights RGB image , we first generate an intensity map I=Imax using, and input the I-map with train able density-k into HVT to generate the HV- map as [13]. Next , HVI concatenate two maps to low-light HVI-map IHVI. I-map and HV- map embedded by two different Conv 3×3 layers and output the I-feature and HV- Feature with shape H×W×C.

The second step involves the utilization of two specific features as inputs to a CID. It outputs the light-up I-feature and denoised HV-feature , both sending to a Conv 3×3 layer in each way and concatenating together to are dual HVI map RHVI. Thereafter, we present the process of inverting HVI to sRGB as our PHVIT . Firstly , it decompose ˆ I toˆ IH,ˆ IV,ˆ II, and clip three maps in the range of[−1,1]. Below we get the density-k of the current iteration, and calculate C .[13]

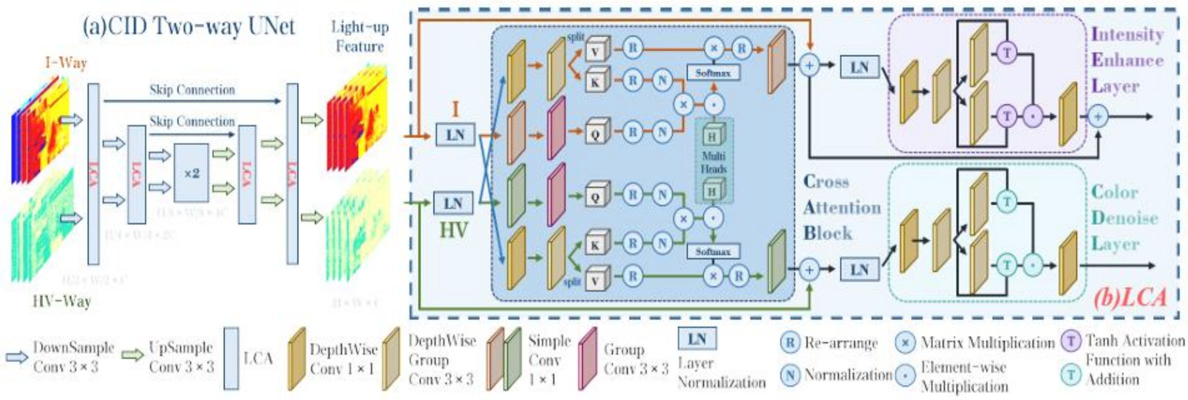


Figure 2.2 – Architecture of CID Two-way UNet.

- In figure 2.2 the core module of Cross-Attention block for LLIE is (b)Lightweight Cross-Attention (LCA), which incorporate a Cross Attention Block (CAB), a Intensity Enhance Layer (IEL) and a Color Denoise Layer (CDL)

2.4.2 Structure

(a) CID Two-way UNet

The CID Two-way UNet embed six LCA modules into a network with two U-shape nets. It contains an encoder with three Down Sample Conv3 × 3 and a decoder with three Up Sample Conv3×3. It inputs the I-way and HV way feature with the shape H × W ×C. Each Down Sample reduces H and W by double and doubles the channels. Each Up Sample Doubles H and W and reduces the channels by double. Finally, it outputs a light-up residual I-way and HV way feature as same as input tensor shape.[12]

(b) Lightweight Cross-Attention

For the purpose of reduce the computational overhead of the Two-way UNet, we specially designed the Lightweight Cross-Attention (LCA) module which only has linear complexity and well-controlled color features based on Transformer.

2.5 LIMITATIONS

1. Dependency on Handcrafted Priors: The method may still have some dependency on handcrafted priors, which could limit its adaptability to various lighting conditions. Handcrafted priors may not be adaptive enough due to the diverse natural scenes and light conditions, potentially impacting the method's performance in real-world scenarios.

2. Computational Load: The proposed method may have limitations in terms of computational loads. While it is mentioned that the method exhibits relatively small parameter and computational loads,[9] it is essential to consider the computational resources required for implementing CIDNet, especially in resource-constrained environments.

3. Color Artifacts and Brightness Artifacts: The proposed HVI color space and CIDNet method may have limitations in handling color artifacts and brightness artifacts in the enhanced images. This suggests that there may be instances where the method does not completely eliminate color and brightness artifacts,[9] potentially impacting the visual quality of the output.

4. Adaptability to Varying Lighting Conditions: While the proposed HVI color space and CIDNet method are designed to adapt to low-light images in different illumination ranges, it is important to consider the method's performance in handling complex and varying lighting conditions.[9] The adaptability of CIDNet to diverse lighting scenarios should be thoroughly evaluated to ensure its effectiveness across a wide range of real-world low-light environments.

CHAPTER 3

Design Methodology

Design Methodology for Low-Light Image Enhancement (LLIE) aims to improve the quality of low-illumination images. It addresses challenges faced by existing methods, such as uncertainty in restoration from diverse brightness degradations and loss of texture and color information. The proposed approach reframes LLIE as learning an image-to-code mapping from low-light images to a discrete codebook, integrating prior knowledge and controllable information transfer to enhance LLIE performance. The approach demonstrates superior robustness to various degradation and outperforms existing state-of-the-art LLIE methods in terms of light enhancement, texture maintenance, noise suppression, and color restoration.

3.1 System Requirement

3.1.1 Software Requirements

Visual Studio Code
Windows 10 (Min)
PyTorch Version - 1.12.0
Python Version - 3.7
HTML
CSS
JavaScript
Cuda Version – 11.7

3.1.2 Hardware Environment Specifications

Hardware that is used in the development of this project is as follows :-

NVDIA RTX 2080 TI – GPU
RAM 8 GB (Min)
CPU Speed – 3.00 GHz (Min)
SSD – 1TB
Processor – i7125OH
Basic Camera

3.2 Proposed Methods

3.2.1 Retinex Model with Paired Low-light Images

The key idea of our method is to learn adaptive priors from low-light image pairs. As a result, our solution needs fewer handcrafted priors and the network is more robust. Note that image pairs are only used in the training phase.

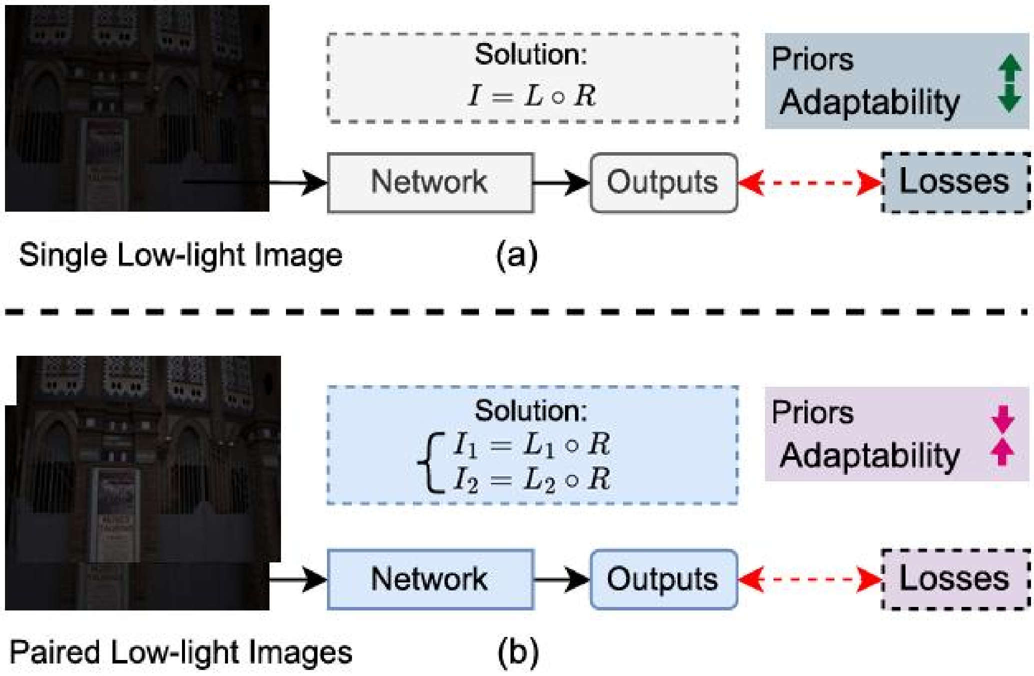


Figure 3.2.1 Comparison between the previous solution (a) and the proposed method (b) from the aspect of Retinex theory.

In Figure 3.2.1 (b) likely utilizes Retinex theory, isolating illumination and reflectance for color and brightness enhancement. It adapts locally for natural-looking results, but may be computationally intensive. In contrast, (a) relies on traditional methods, possibly lacking in perceptual quality and adaptability to varying lighting conditions compared to the Retinex-based approach.

According to the Retinex theory, low-light image I can be decomposed into illumination L and reflectance R as:

$I = L \circ R,$

(1)

where $\circ$ denotes the element-wise multiplication. Illumination L describes the light intensity of objects. L should be piece-wise continuous and texture less. Reflectance R represents the physical properties of the objects. R should contain textures and details in the observed image. The Retinex decomposition is highly ill-posed. Various methods have been proposed to handle this problem. A generic solution of the Retinex decomposition is to minimize the following energy function:

$$E(I,L,R) = \lambda_1 \quad E\,data \quad (I,L,R) + \lambda_2 \quad E\,smooth \quad (L) + \lambda_3 \quad Etex \quad (R)$$

Where:

- I is the observed low-light image.
- L is the estimated illumination component.
- R is the estimated reflectance component.
- Edata(I,L,R) represents the data fidelity term, ensuring that the reconstructed image closely matches the observed low-light image.
- Esmooth(L) is a smoothness term on the illumination component, encouraging piece-wise continuity and texturelessness.
- Etex(R) is a texture preservation term on the reflectance component, promoting the retention of textures and details.
- $\lambda 1, \lambda 2, \lambda 3$ are trade-off parameters controlling the importance of each term.

However, handcrafted priors are commonly not adaptive enough due to the diverse natural scenes and light conditions. In this paper, instead of exploiting handcrafted priors for L and R from a single image, we apply paired lowlight images to automatically learn adaptive priors in a data driven fashion. Those low-light image pairs share the same scene content but different illumination. Mathematically, Retinex decomposition with low-light image pairs can be expressed as:

$$I1 \quad = L1 \quad \circ \quad R1$$

Where:

- I1 and I2 are two low-light images from the paired dataset, sharing the same scene content but having different illumination conditions.

- L1 and L2 represent the estimated illumination components corresponding to I1 and I2 respectively.
- R1 and R2 denote the estimated reflectance components corresponding to I1 and I2 respectively.

3.2.2 Network Structure

The whole pipeline of our method is illustrated in Fig. We use L-Net and R-Net to estimate the illumination and reflectance components, respectively. L-Net and R-Net are very similar and simple, both of which contain five convolutional layers. The activation function of the first four convolutional layers is Re LU. L-Net and R-Net end with a sigmoid layer to normalize the output into [0, 1]. According to the Retinex theory, the three-color channels are assumed to have the same illumination. Therefore, the output channel of L-Net is set as 1 while that of R-Net is set as 3. Note that, this paper does not focus on designing modernistic network structures. In contrast, we aim at providing a generic solution for LIE under paired low-light instances. In our experiments, we found that simple network already achieves comparable performance. Apart from L-Net and R-Net, we introduce P-Net to remove inappropriate features from the original image. Specifically, the structure of the P-Net is identical to the R-Net. In the training phase, the original low-light image pairs I1 and I2 are first taken into the P-Net, yielding two optimized versions i1 and i2. Then, L-Net and R-Net are applied to estimate the latent illumination (L1 and L2) and reflectance (R1 and R2). To optimize the network, three loss functions are designed in PairLIE. The first is the projection loss LP that measures the difference between I and i. The second is the reflectance consistency loss LC calculated based on R1 and R2. The third is the Retinex loss LR that restricts the decomposed components to satisfy the Retinex theory. In the testing period, given a low-light image, P-Net, R-Net, and L-Net are applied to calculate the final enhanced image:

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

Where λ represents the illumination correction factor, and I_{en} denotes the enhanced image.

3.2.3 Projection Loss

Instead of performing the deep Retinex decomposition on the original low-light image, we propose to first remove inappropriate features to ensure the input can be accurately decomposed by an ideal Retinex model.

$$\mathrm{L}p = \| I1 - i1 \|_2^2$$

where i1 refers to the projected image. The loss function Lp transforms the original image to a specific image that is more suitable for the Retinex decomposition. Concretely, this image is expected to be noise free since we do not consider the noise component in the Retinex model.[9] Besides, some useless features are also discarded in this stage. Our projection loss can be interpreted from the view of error reallocation, which can be expressed as:

It's crucial to note that the projection loss must work in conjunction with other constraints to prevent trivial solutions, such as 1= 1.

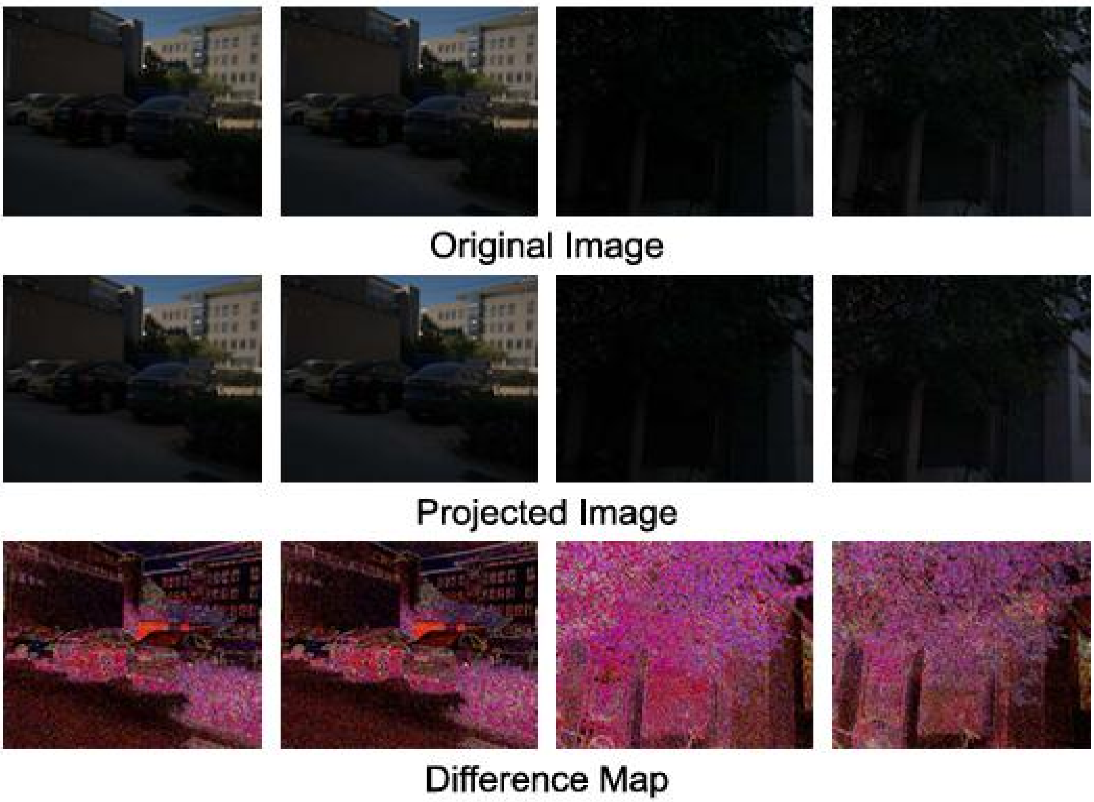


Figure 3.2.2 Example of projected images and corresponding difference maps.

Figure 3.2.2 displays projected images alongside their corresponding difference maps. These maps highlight variations between the original and projected images, aiding in visualizing the effectiveness of the projection process.[9] Such comparisons are crucial for evaluating the accuracy and fidelity of image projection techniques.

3.2.4 Reflectance Consistency Loss

Reflectance consistency loss LC is calculated based on low-light image pairs and the Retinex theory. Compared with handcrafted priors, LC is more accurate and adaptive because it reveals the physical properties of the objects. Mathematically, LC is formulated as:

$$LC = \| R1 - R2 \|_2^2$$

Where:

- LC represents the reflectance consistency loss.
- R1 and R2 are the reflectance components of paired low-light images.

The loss function essentially penalizes the difference between the reflectance components of two paired low-light images. By minimizing this loss, the network is encouraged to predict similar reflectance components for corresponding objects in the low-light image pairs, thereby enhancing the consistency of the reflectance across the images.

Since sensor noise hidden in dark regions will be amplified when the contrast is improved. To cope with the noise issue, existing approaches either add a smoothness term on the estimated reflectance or perform a denoise operation after the enhancement. In our method, the sensor noise can be implicitly removed.

This is because the two lowlight images contain independent noise of the same scene. As discussed in [5], paired noisy images can be leveraged to train denoising models. This is because noise is random and different in two images, the deep network cannot fit the noise in one image to another. In our case, the two lowlight images can help each other to remove noise during the Retinex decomposition. Therefore, PairLIE does not require additional handcrafted constraints of noise.[6]

3.2.5 Retinex Loss

Some basic constraints of Retinex decomposition are introduced, which can be formulated as:

$LR = \| R \circ L - i \|_2^2 + \| \| R - \text{stopgrad}(L) i \| \|_2^2 + \| L - L0 \|_2^2 + \| \nabla L \|$

Where:

- LR represents the Retinex loss.
- R and L are the decomposed reflectance and illumination components, respectively.[6]
- i denotes the input image.
- $L0$ is the initial estimation of the illumination.
- ∇ represents the horizontal and vertical gradients.

This formulation incorporates several constraints:

1. The first term $\|R \circ L - i\|_2^2$ ensures that the product of the decomposed components R and L is close to the input image i, facilitating accurate reconstruction.
2. The second term $\|R - i\text{stopgrad}(L)\|_2^2$ enforces consistency between the reconstructed image R and the input image ii based on the estimated illumination L.
3. The third term $\|L - L0\|_2^2$ encourages the estimated illumination to be close to the initial estimation $L0$
4. The fourth term $\|\nabla L\|$ imposes a smoothness constraint on the illumination by penalizing its gradients.

Overall, these constraints guide the Retinex decomposition process to ensure that the decomposed components effectively capture the essential features of the input image while maintaining consistency and smoothness[3].

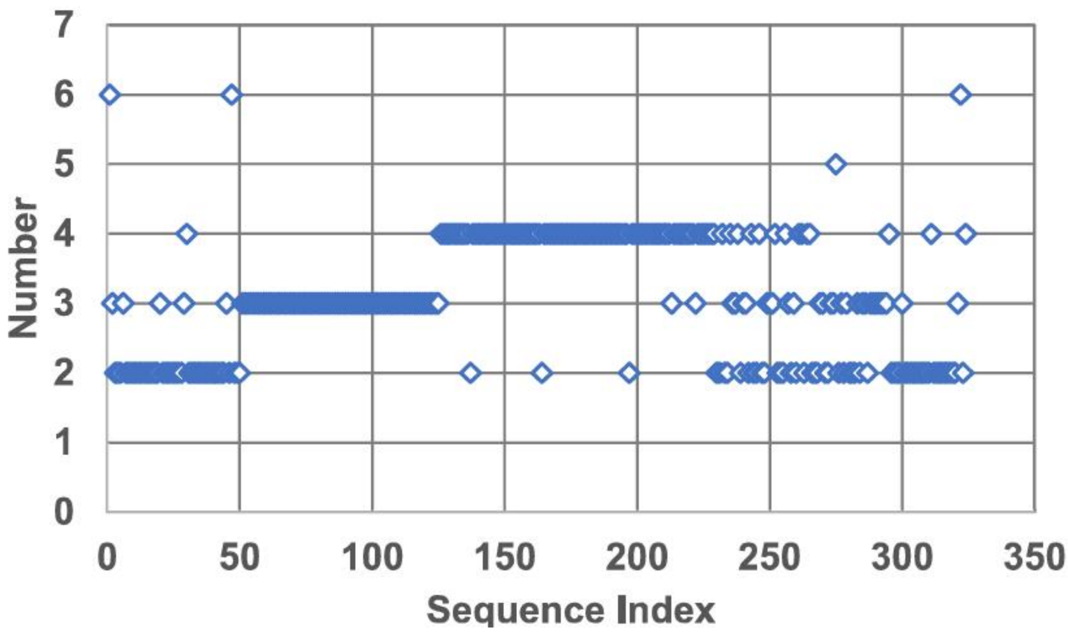


Figure 3.2.3 Statistics

In Fig 3.2.3 ,Statistics concerning sequence length of the collected data. Most sequences have 2 to 4 low-light images.

3.2.6 Overall Loss

The overall loss function LAll for training PairLIE is a linear combination of three individual losses, namely LP, LC , and LR. It is expressed as:

$L_{\mathrm{All}}=w0_{LP}+w1_{LC}+w2_{\mathrm{LR}}$

Where:

- $w0$, $w1$, and $w2$ are the weights assigned to each loss term.

This formulation allows adjusting the importance of each loss term during training by controlling the respective weights $w0$, $w1$, and $w2$. By tuning these weights, the training process can prioritize different aspects of the model's performance, such as perceptual quality, reflectance consistency, and adherence to Retinex constraints.

3.2.7 Datasets and Criteria

As described before, low-light image pairs collected from SICE and LOL are applied to train PairLIE. We select another 50 sequences (150 images) from SICE and use the official evaluation set (15 images) of LOL to measure the model performance. Since SICE and LOL contains reference images, we employ PSNR, SSIM, LPIPS, and

DeltaE with CIE2000 standard to objectively evaluate the performance of each method. A higher PSNR/SSIM score indicates the result is closer to the reference. A lower LPIPS/DeltaE value denotes better enhancement performance. Furthermore, we adopt the MEF dataset for visual comparisons.

3.2.8 Training Data Collection

We collect low-light image pairs from SICE (part2) and LOL (training set) , which contain multi-exposure images. These datasets adopt some specific operators to deal with the misalignment caused by camera shaking or object moving. Note that the SICE dataset consists of both under and over-exposed images. We only select underexposure and well-aligned cases for constructing the training set. As a result, we collect 324 sequences (a total of 1000 low-light images). As shown, each sequence includes 2 to 6 samples. In the training phase, we randomly select two images from each sequence to constitute a pair.

IMPLEMENTATION, RESULTS AND ANLYSIS

We first describe implementation details, evaluation datasets, and Result. Then, we present the quantitative and qualitative comparisons with state-of-the-heart methods. Finally, we conduct ablative experiments to validate each component.

4.1 Implementation

We implement the method with PyTorch. In the training phase, we randomly crop images to the size of 128 × 128. A batch size of 1 is applied. We use ADAM with the initial learning rate of 1×10^{-4} to optimize the network. The number of training epochs is set to 400 and 3000. The learning rate is half-decayed per 100 epochs. The default correction factor λ is 0.2. In extremely dark cases such as the LOL datasets, we set λ = 0.14. As for the hyper-parameter w0, w1, and w2 we set w0 = 500, w1 = w2 = 1, empirically.

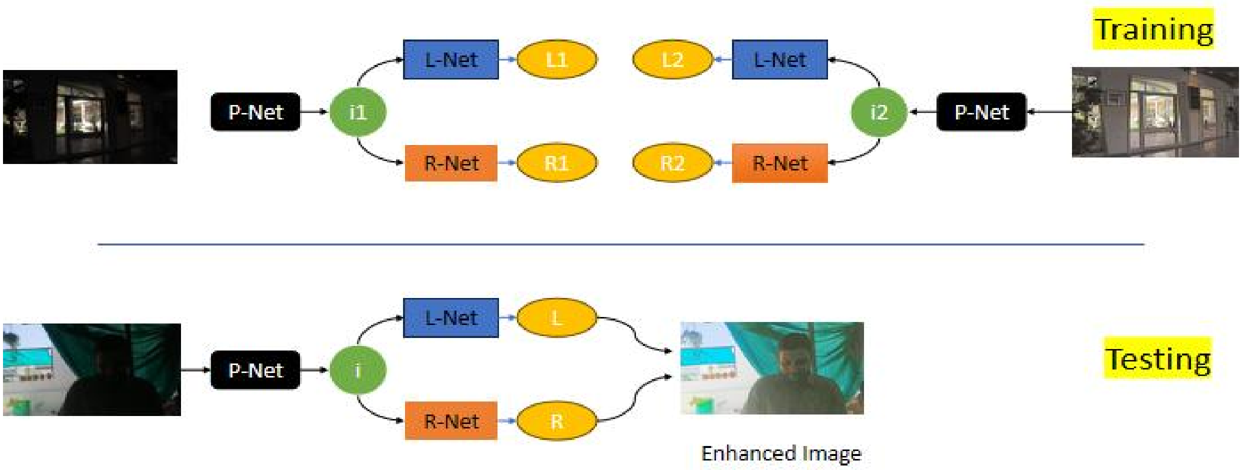


Figure 4.1 System Architecture

Fig 4.1 Describe the Model Architecture and the component used :

- **L-Net:** A convolutional neural network (CNN) designed to estimate the illumination component of an image. It consists of five convolutional layers, with ReLU activations for the first four layers and a sigmoid activation for the output layer to normalize illumination values between 0 and 1.
- **R-Net:** Similar in structure to L-Net but designed to estimate the reflectance component. It has three output channels corresponding to the RGB color channels.

- **P-Net:** An auxiliary network identical to R-Net, used to preprocess images by removing inappropriate features before Retinex decomposition.

In fig 4.1 HVI Color Transform (HVIT): It takes an sRGB image as input and generates HV feature, Intensity feature, and HVI map as outputs. CID Two-way UNet: This module performs the main processing, utilizing a Two-way UNet architecture. Lastly, we apply Perceptual-inverse HVI Transform (PHVIT), which takes a light-up HVI map as input and transforms it to an sRGB enhanced image.

Retinex Decomposition:

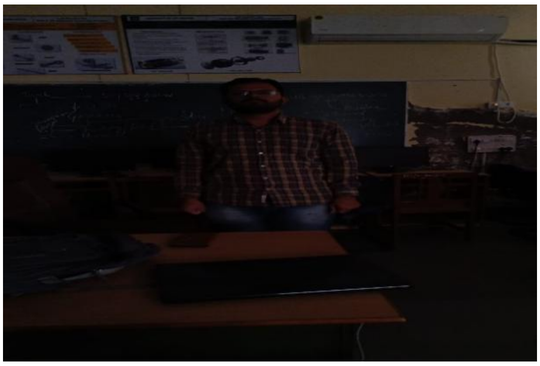
- **Illumination (L) and Reflectance (R):** According to the Retinex theory, a low-light image I can be decomposed into illumination and reflectance such that $=L \circ R$, where $\circ$ denotes element-wise multiplication.

Loss Functions:

- **Projection Loss (L*P*):** Ensures that the input image is transformed into a form suitable for Retinex decomposition, reducing noise and irrelevant features.
- **Reflectance Consistency Loss (L*C*):** Enforces consistency between the reflectance components of paired low-light images, leveraging their shared scene content.
- **Retinex Loss (L*R*):** Incorporates basic constraints of Retinex decomposition to ensure accurate reconstruction and stabilization of the network training.

4.2 Result

INPUTS **OUTPUTS**



4.3Analysis

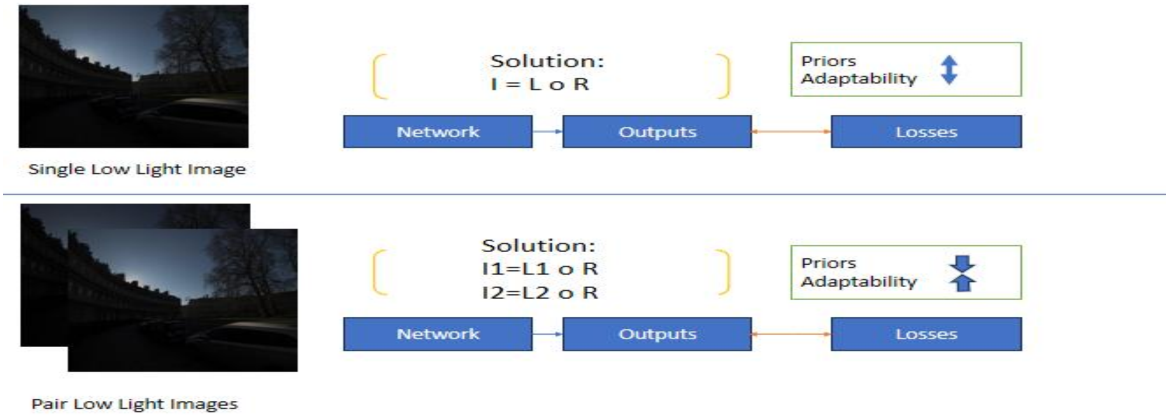


Fig 4.2 Comparison Between previous solution and proposed method

Fig 4.2 , Shows that our solution needs fewer handcrafted priors and the network is more robust. Note that image pairs are only used in the training phase.

Quantitative Metrics:

- **PSNR (Peak Signal-to-Noise Ratio):** Measures the reconstruction quality.
- **SSIM (Structural Similarity Index):** Assesses image similarity based on structural information.

Method	LOL				SICE			
	PSNR	SSIM	LPIPS,	DeltaE t	PSNR	SSIM1	LPIPS.	DeltaE t
SDD [9]	13.34	0.637	0.743	21.83	15.35	0.741	0.232	16.08
STAR [34]	12.91	0.518	0.366	23.46	15.17	0.727	0.246	16.35
MBLLEN [20]	17.86	0.727	0.225	13.68	13.64	0.632	0.297	18.60
RetinexNet [32]	17.55	0.648	0.379	12.69	19.89	0.783	0.276	8.715
GLADNet [30]	19.72	0.680	0.321	12.28	19.98	0.837	0.203	8.947
KinD [38]	17.65	0.775	0.171	12.49	21.10	0.838	0.195	8.009
DRBN [36]	16.29	0.551	0.260	13.44	15.58	0.522	0.289	13.78
URetinexNet [33]	19.84	0.826	0.128	10.65	21.64	0.843	0.192	7.728
ZeroDCE [7]	14.86	0.559	0.335	18.81	18.69	0.810	0.207	11.93
RRDNet [41]	11.40	0.457	0.362	26.43	13.28	0.678	0.221	19.64
RUAS [18]	16.40	0.500	0.270	16.85	13.18	0.734	0.363	16.81
SCI [22]	14.78	0.522	0.339	19.52	15.95	0.787	0.235	13.71
EnlightenGAN [12]	17.48	0.651	0.322	14.50	18.73	0.822	0.216	10.42
Ours	**19.51**	**0.736**	**0.248**	**10.80**	**21.32**	**0.840**	**0.216**	**7.835**

Qualitative Analysis:

- **Visual Comparisons:** Compare enhanced images from our method with those from state-of-the-art methods to assess improvements in brightness, color, contrast, and naturalness.

- **Noise Suppression:** Evaluate the effectiveness of noise suppression by examining enhanced images for noise artifacts.

Fig 4.3 Qualitative Analysis

Fig 4.3, Example of projected images and corresponding difference maps. A smaller value indicates a higher similarity between the projected and original images.

CHAPTER 5
FUTURE WORK

Future work in the field of low-light image enhancement could focus on addressing the limitations and challenges identified in the existing methods. Some potential areas for future research and development include:

1. Noise Artifact Elimination: Future work could aim to develop more effective methods for eliminating noise artifacts in low-light image enhancement. This could involve exploring advanced noise reduction techniques and incorporating them into the enhancement process to improve the overall visual quality of the enhanced images.

2. Accurate Color Mapping: There is a need for future research to focus on achieving more accurate color mappings in low-light image enhancement. This could involve developing innovative approaches to ensure that the colors in the enhanced images are faithfully represented, even in challenging low-light conditions.

3. Adaptability to Varying Lighting Conditions: Future work could explore methods that enhance the adaptability of low-light image enhancement techniques to handle complex and varying lighting conditions. This could involve developing algorithms that can effectively adjust to different illumination scales and color variations in diverse low-light environments.

4. Computational Efficiency: There is a potential for future research to focus on improving the computational efficiency of low-light image enhancement methods. This could involve exploring techniques to reduce parameter loads and computational loads while maintaining high-quality enhancement results.

5. Evaluation Metrics: Future work could involve the development of comprehensive evaluation metrics specifically tailored to assess the performance of low-light image enhancement methods. This could help in objectively comparing different approaches and identifying their strengths and weaknesses.

6. Real-world Applications: Further research could focus on the practical application of low-light image enhancement techniques in real-world scenarios, such as surveillance, autonomous driving, and medical imaging. This could involve conducting extensive validation studies and field tests to assess the effectiveness of the methods in real-world settings.

Overall, future work in low-light image enhancement should aim to address the current limitations, improve the robustness and adaptability of the methods, and facilitate their practical implementation in various domains.

CONCLUSIONS

This method is a reference-free approach for low-light image enhancement that benefits from both Retinex-based and learning-based solutions. By learning adaptive constraints from low-light image pairs, this reduces the dependence on handcrafted priors and thus generalizes well on various scenes. To assist the decomposition, this first removes inappropriate features in the original image and then implements the decomposition on the optimized image. Extensive experiments on public benchmarks show that this outperforms the state-of-the-art unsupervised methods significantly. In future works, we will concentrate on exploiting effective priors of illumination in a data-driven fashion.

REFERENCES

[1] Yixu Feng, Cheng Zhang, Pei Wang, Peng Wu, Qingsen Yan, Yanning Zhang. “You Only Need One Color Space: An Efficient Network for Low-light Image Enhancement”. In CVPR 8 Feb 2024

[2] A. Brateanu, R. Balmez, A. Avram, C. Orhei. “LYT-Net: Lightweight YUV Transformer-based Network for Low-Light Image Enhancement”. In CVPR 24 Feb 2024

[3] Hai Jiang, Ao Luo, Songchen Han, Haoqiang Fan, Shuaicheng Liu. “ Low Light image enhancement with Wavelet-based Diffusion Models”. In CVPR 1 Jun 2023

[4] Zhen Tian ,, Peixin Qu , Jielin Li , Yukun Sun , Guohou Li , Zheng Liang and Weidong Zhang . "A Survey of Deep Learning-Based Low-Light Image Enhancement" . In MDPI 10 August 2023

[5] Feng Zhang, Yuanjie Shao, Yishi Sun, Kai Zhu, Changxin Gao, Nong Sang . “Unsupervised Low-Light Image Enhancement via Histogram Equalization Prior” . in CVPR 3 Dec 2021f++

[6] Tao Huang, Songjiang Li, Xu Jia, Huchuan Lu, and Jianzhuang Liu. Neighbor2neighbor: Self-supervised denoising from single noisy images. In CVPR, pages 14781– 14790, 2021.

[7] Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, Runmin Cong. “Zero- Reference Deep Curve Estimation for Low-Light Image Enhancement” . In CVPR 22 Mrch 2020

[8] Hunsang Lee, Kwanghoon Sohn, and Dongbo Min. Unsupervised low-light image enhancement using bright channel prior. IEEE SPL, 27:251–255, 2020.

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

[10] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2Noise: Learning image restoration without clean data. In ICML, pages 2965–2974, 2018.

[11] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In CVPR, pages 3291–3300, 2018.

[12] Xiaojie Guo, Yu Li, and Haibin Ling. Lime: Low-light image enhancement via illumination map estimation. IEEE TIP, 26(2):982–993, 2017.

[13] Xueyang Fu, Delu Zeng, Yue Huang, Yinghao Liao, Xinghao Ding, and John Paisley. A fusion-based enhancing method for weakly illuminated images. Signal Processing, 129:82–96, 2016.