Heliyon 9 (2023) e14558

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Review article

# A survey on image enhancement for Low-light images

Jiawei Guo [a,b], Jieming Ma [b,*], Ángel F. García-Fernández [c,d], Yungang Zhang [e], Haining Liang [b]

[a] Department of Computer Science, University of Liverpool, Liverpool, UK
[b] School of Advanced Technology, Xi'an Jiaotong-Liverpool University (XJTLU), Suzhou, China
[c] Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, UK
[d] ARIES research center, Universidad Antonio de Nebrija, Madrid, Spain
[e] School of Information Science Yunnan Normal University, Kunming, China

ARTICLE INFO

*Keywords:*
Image enhancement
Low-light images
Image processing
Deep learning

ABSTRACT

In real scenes, due to the problems of low light and unsuitable views, the images often exhibit a variety of degradations, such as low contrast, color distortion, and noise. These degradations affect not only visual effects but also computer vision tasks. This paper focuses on the combination of traditional algorithms and machine learning algorithms in the field of image enhancement. The traditional methods, including their principles and improvements, are introduced from three categories: gray level transformation, histogram equalization, and Retinex methods. Machine learning based algorithms are not only divided into end-to-end learning and unpaired learning, but also concluded to decomposition-based learning and fusion based learning based on the applied image processing strategies. Finally, the involved methods are comprehensively compared by multiple image quality assessment methods, including mean square error, natural image quality evaluator, structural similarity, peak signal to noise ratio, etc.

## 1. Introduction

Computer vision technology and deep learning are more and more widely used in many fields, such as medical image processing [1], automatic driving [2], face recognition [3], object detection [4]. However, due to physical constraints, such as weak illumination, limited exposure time and unsuitable camera angles, images often suffer multiple degradations, including but not limited to poor visibility, low contrast, backlight, shadow, and nighttime. Examples of such pictures are shown in Figs. 1(a) to 1(d). Because of non-uniform illumination and low contrast, information about the image is masked or lost properly, which restricts utilization for real world applications such as applications of remote sensing images [5], lane detection [6], etc.

Image enhancement is one of the main tasks of image processing, which aims to make images match the visual response characteristics and selectively highlight the features of interest in images by adding some information or transforming data to original images by certain methods. The main purposes of image enhancement are to expand the difference between the features of different objects in images, suppress the features that are not interested, improve the image quality, enrich the amount of information, strengthen the image interpretation and recognition effect, and meet the requirements of some special analysis.

* Corresponding author.
*E-mail address:* Jieming.Ma@xjtlu.edu.cn (J. Ma).

https://doi.org/10.1016/j.heliyon.2023.e14558
Received 20 November 2022; Received in revised form 22 January 2023; Accepted 9 March 2023

Available online 16 March 2023

(a) Nighttime (b) Shadow (c) Indoor (d) Backlight

**Fig. 1.** Degradations of images.

The purpose of this paper is to give a comprehensive literature review of image enhancement from the perspective of algorithms. There are several existing reviews on image enhancement. Wencheng et al. [7] review the main techniques of low-light image enhancement developed over the past decades. Rasheed et al. [8] review Retinex-based low-light enhancement methods, and compares them with other state-of-the-art low-light enhancement methods. Chongyi et al. [9] introduce low-light image and video enhancement algorithms based on deep learning. Sobbahi et al. [10] review the application of image enhancement to object classification and recognition tasks. However, these works are either simply linear introductions to image enhancement, or only focus on the structures and applications of methods while ignoring the derivation of image digital theory and the connections between the methods. This paper first gives a comprehensive review of traditional learning methods. In the section of machine learning-based methods, this paper not only simply divides the algorithms into end-to-end learning and unpaired learning categories, but also innovatively summarizes decompression-based and fusion-based algorithms which are highly integrated with traditional algorithms. The main contributions of this paper are the following:

1. In this paper, a comprehensive study of the traditional image enhancement algorithms is discussed, which will help the readers understand the advantages and shortages of image enhancement methods from a mathematical perspective. The algorithmic view of traditional enhancement methods is to design appropriate filtering techniques and modify pixel values based on some operations on images, which will improve the quality of the images that are apparent visual to humans. This paper divides the traditional enhancement methods into three categories: (1) gray level transformation methods; (2) histogram equalization methods; (3) Retinex-based methods.
2. Unlike traditional algorithms, machine learning-based algorithms enhance low-light images by learning image features under normal lighting conditions, which requires a large amount of data for training. For machine learning-based enhancement methods, this paper focuses on the mathematical theory contained in the algorithm and divides them into four categories: (1) end-to-end learning methods; (2) decomposition-based learning methods; (3) fusion-based learning methods; (4) unpaired learning methods.
3. In order to analyze the performance of different image enhancement algorithms from both subjective and objective perspectives, we reproduce the traditional methods and 11 machine learning algorithms mentioned in this paper, and quantitatively analyze and compare these algorithms based on the image quality assessment methods.

The remainder of this paper is organized as shown in Fig. 2. Section 2 introduces traditional image enhancement algorithms. In Section 3, novel machine learning-based algorithms are introduced, including common loss functions and the application of image enhancement in high-level visual tasks. Some commonly used datasets and evaluation methods, and the experiment results of some representative machine learning-based image enhancement algorithms are shown in Section 4. Section 5 summarizes the conclusions and gives several suggestions for future research directions.

## 2. Traditional methods

The gray value of the dark area of an image is usually small, so the idea of traditional image enhancement algorithms is to design a mathematical formula or a filtering method to adjust the gray value of the image. In this paper, the traditional image enhancement algorithms are divided into three categories according to different enhancement ideas: (1) gray level transformation methods; (2) histogram equalization methods; (3) Retinex-based methods.

### *2.1. Gray level transformation methods*

In the case of underexposure or over-exposure, the grayscale value of images may be limited to a relatively small range, which may lead to the problems of blurring and gray level disappearing. Gray level transformation is an essential method of image enhancement used to improve the image display effect, belonging to the spatial domain processing method [11,12]. It can increase the dynamic range of the image and expand the image contrast so that the image can be clear. The essence of gray level transformation is to modify the grayscale of each pixel of the image according to certain rules to change the grayscale range of the image. Typical gray level transformation methods can be divided into linear and nonlinear transformations.

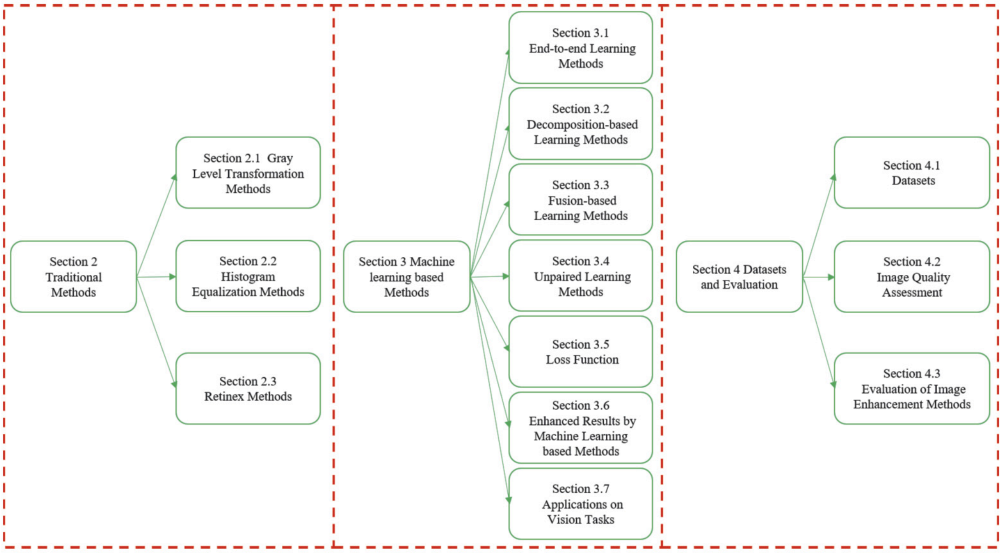



**Fig. 2.** Section Organization.

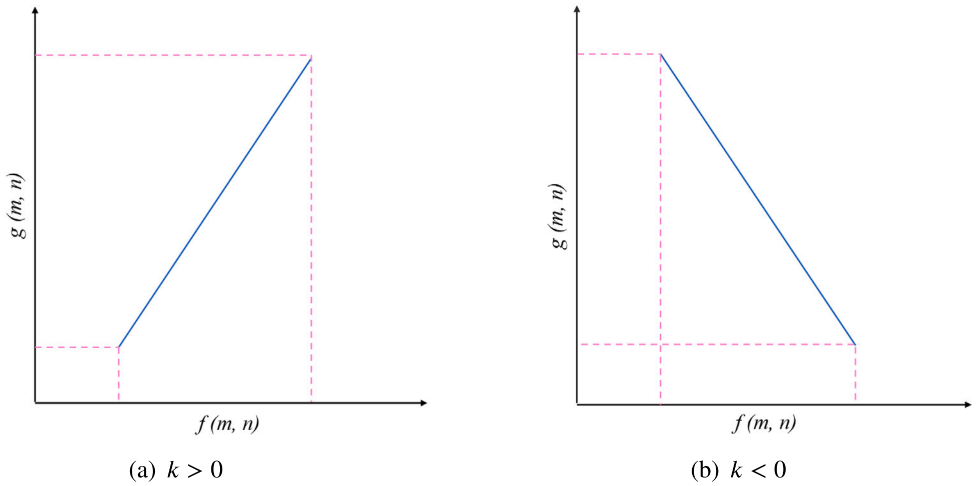



**Fig. 3.** Linear gray level transformation.

### 2.1.1. *Linear gray level transformation*

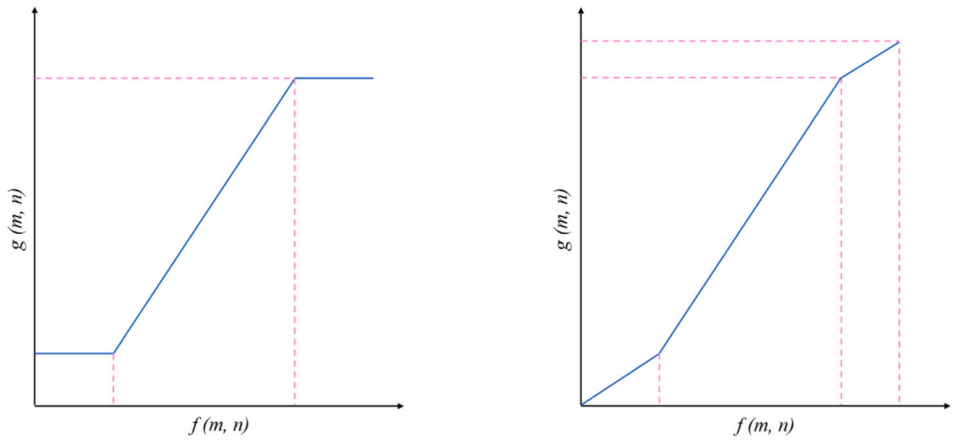
Linear gray level transformation employs linear equations to map gray values. Suppose the gray value of an original image $f(m,n) \in [a,b]$ and the gray value after linear transformation $g(m,n) \in [c,d]$, the formula of linear gray level transformation is as Eq. (1):

$$g(m,n) = k[f(m,n) - a] + c, \tag{1}$$

in which, $(m,n)$ determines the coordinates of the pixels, and $k = \frac{d-c}{b-a}$ is the slope of the transformation function. The transformations of $k > 0$ and $k < 0$ are shown in the Figs. 3(a) and 3(b).

According to the values of $[a,b]$ and $[c,d]$, there are four possible situations.

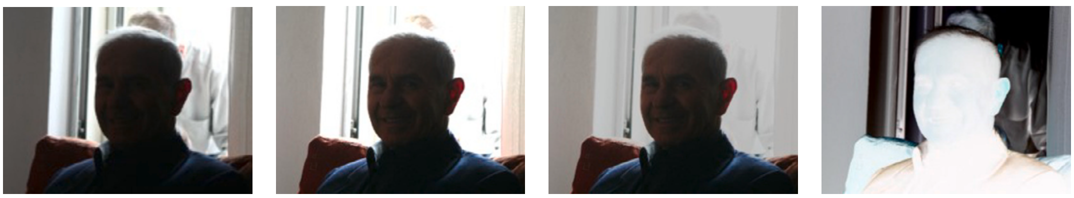
1. **Dynamic Range Extension.** If $[a,b] \subset [c,d]$, i.e., $k > 1$, the size of the dynamic range of gray values will be expanded, and the dynamic range of image display equipment can be fully utilized, thus improving the problem of insufficient image exposure.
2. **Change of Gray Value Range.** If $[a,b] = [c,d]$, i.e., $k = 1$, the size of the gray value range of the transformed image will not change. However, the gray value range will shift with the values of a and c.
3. **Dynamic Range Reduction.** If $[a,b] \supset [c,d]$, i.e., $k < 1$, the size of dynamic range of image gray value will be reduced.
4. **Reverse.** If $k < 0$, The gray values of the transformed image will be reversed. The parts of the transformed image will darken where they were initially bright, and the parts of the transformed image will become bright where they were initially dark.




(a) Expand the interesting interval and make the others static.

(b) Expand the interesting interval and suppress the others.

**Fig. 4.** Piece-wise gray level transformation.



(a) Original image (b) Linear transformation (c) Piece-wise transformation (d) Reverse transformation

**Fig. 5.** Enhanced results by gray transformation methods.

### *2.1.2. Piece-wise gray level transformation*

Sometimes, the gray values of the whole image do not need to be adjusted. Only the values in some areas of interest should be stretched or compressed. Based on the application scenarios, the piece-wise linear gray level transformation can be divided into two situations.

1. **Expand the interesting interval and make the others static.** For the interesting interval $[a, b]$, the linear transformation with a slope greater than 1 is used to expand the gray value, and the other interval is expressed as $a$ or $b$, as shown in Eq. (2).

$$g(m,n) = \begin{cases} a & f(m,n) < a \\ c + \frac{d-c}{b-a}[f(m,n) - a] & a \le f(m,n) \le b \\ b & f(m,n) < b \end{cases} . \tag{2}$$

2. **Expand the interesting interval and suppress the others.** For the interesting interval, $[a, b]$, the linear transformation with a slope greater than 1 expands the gray value. For the other interval, the gray value is suppressed by the linear transformation with a slope less than 1, as shown in Eq. (3).

$$g(m,n) = \begin{cases} \frac{c}{a} f(m,n) & f(m,n) < a \\ c + \frac{d-c}{b-a}[f(m,n) - a] & a \le f(m,n) \le b \\ d + \frac{N-d}{M-b}[f(m,n) - b] & b \le f(m,n) \le M \end{cases} . \tag{3}$$

The two kinds of piece-wise gray level transformation are shown in Figs. 4(a) and 4(b).

Figs. 5(a) to 5(d) show the original image and the transferred images by the linear transformation, piece-wise transformation, reverse transformation, respectively. Linear transformation enhances the whole picture, so the bright background becomes brighter. Piece-wise transformation enhances the different areas according to the split points, and reverse transformation reverses the gray values of the image.

### *2.1.3. Logarithmic transformation*

The general formula for the logarithmic transformation is as shown in Eq. (4)

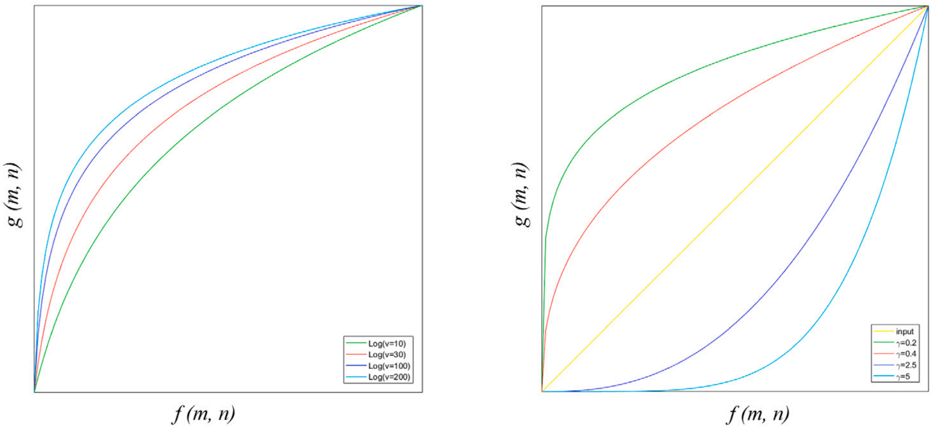
$$g(m,n) = \lambda \log_{v+1}(1 + v \cdot f(m,n)), \tag{4}$$

(a) Original image (b) $log = 10$ (c) $log = 30$ (d) $log = 100$ (e) $log = 200$

**Fig. 6.** Enhanced results by the logarithmic transformation.

(a) Original image (b) $gamma = 0.2$ (c) $gamma = 0.4$ (d) $gamma = 2.5$ (e) $gamma = 5$

**Fig. 7.** Enhanced results by the gamma transformation.




(a) Logarithmic transformation (b) Gamma transformation

**Fig. 8.** The functions of Logarithmic transformation and Gamma transformation.

where $\lambda$ is an adjustment constant, which is used to adjust the gray value to make the transformed images conform to the actual requirements, and $\upsilon + 1$ is the base number. Logarithmic transformation of an image means replacing all pixel values present in the image with its logarithmic values. Logarithmic transformation is used for image enhancement as it expands dark pixels of the image and compresses bright pixels. The function of logarithmic transformation is shown in Fig. 8(a). Figs. 6(a) to 6(e)show the enhanced images by the logarithmic transformation under different parameters. As can be seen from the image, the brightness of the image increases with the parameter $\upsilon$, and the dark area is enhanced faster.

### 2.1.4. Gamma transformation

The general formula for the Gamma transformation is as Eq. (5).

$$g(m,n) = \lambda(f(m,n) + \varepsilon)^{\gamma}, \tag{5}$$

in which, $\lambda$ and $\gamma$ are constants. $\mathcal{E}$ is set to avoid the situation that the base is 0. When $\gamma < 1$, the gray value of the image will map to the high brightness. On the contrary, when $\gamma > 1$, the gray value of the image will map to the low brightness. The function of Gamma transformation is shown in Fig. 8(b). Figs. 7(a) to 7(e) show the results enhanced by gamma transformation under different parameters.

## 2.2. Histogram equalization methods (HE)

Suppose the gray histogram of an image almost covers the whole range of gray values, and the distribution of the whole gray values is approximately uniform except for some prominent gray values. In that case, the image has an extended dynamic range

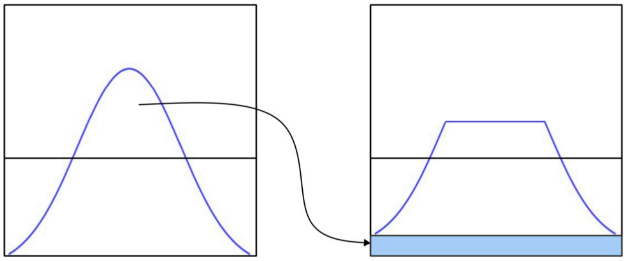


**Fig. 9.** Processing of clipping value of the histogram.

of gray values and a high contrast, and the details of the image could be relatively abundant. In the HE algorithm, cumulative distribution function (CDF) is applied to adjust the output gray level, so that the gray distribution of the image becomes uniform.

### 2.2.1. Global histogram equalization (GHE)

Global histogram equalization, the basic HE algorithm, is processed by considering the whole image, while local histogram equalization equalizes a part of the histogram to enhance more details of the image. In GHE [13–15], the input image is first divided into different pixel levels according to the gray value of the pixel and calculated the histogram of each gray value level in the input image pixels. CDF can be obtained by adding up the histograms. Finally, the gray value after transformation can be calculated using gray transform mapping. The GHE algorithm can be stated as follows.

1. Set gray value level $r_k$, and the number of gray value levels is $l$;
2. Calculate the proportion of each gray value level to the total number of pixels in the original image $P(k)$;
3. Calculate cumulative distribution function (CDF) as Eq. (6):

$$c(k) = \sum_{k=1}^{l} P(k). \tag{6}$$

4. Calculate the transformed gray histogram according to the gray level transformation mapping, and round it to the nearest integer, $INT()$ is rounding to an integer, as shown in (7):

$$Y(k) = INT[(l-1)c(k) + 0.5]. \tag{7}$$

5. Count the number of pixels of the gray level after transformation, and calculate the transformed histogram.

GHE can effectively enhance overall darker or brighter images. However, it is difficult for a global algorithm to enhance the local region of the input image, and it may make parts of the image too bright or too dark.

### 2.2.2. Adaptive histogram equalization (AHE)

Sometimes, GHE cannot meet the actual requirements since it may cause details to disappear in regions that do not need enhancement. The basic idea of AHE [14–16] is to separate an image into several sub-blocks, and each sub-block is processed by histogram equalization, respectively. The AHE algorithm can be stated as follows.

1. Set the size of a window, and select a subblock of the input image according to the window;
2. Apply HE algorithm to the subblock, and record the output;
3. Move the window horizontally or vertically and repeat 1 and 2 until all pixels in the input image are modified;
4. Organize all the enhanced subblocks into one image as output.

### 2.2.3. Contrast limited adaptive histogram equalization (CLAHE)

AHE tends to overamplify the contrast in near-constant regions of the image because of the high concentration of histograms in these areas. As a result, noise may be amplified in near-constant regions. Contrast Limited Adaptive Histogram Equalization (CLAHE) [17] is a kind of adaptive histogram equalization in which the contrast amplification can be limited to reduce the problem of noise amplification.

CLAHE limits the amplification of noise by clipping the histogram at a predefined value which can limit the slope of the CDF. The value at which the histogram is clipped depends on the normalization of the histogram and, thereby, on the neighborhood region's size. The clipping value of the histogram needs to be evenly distributed in the whole gray range, as shown in Fig. 9, to ensure that the total area of the histogram is consistent with that before clipping. Fig. 10 shows the images processed by the three kinds of HE algorithms.

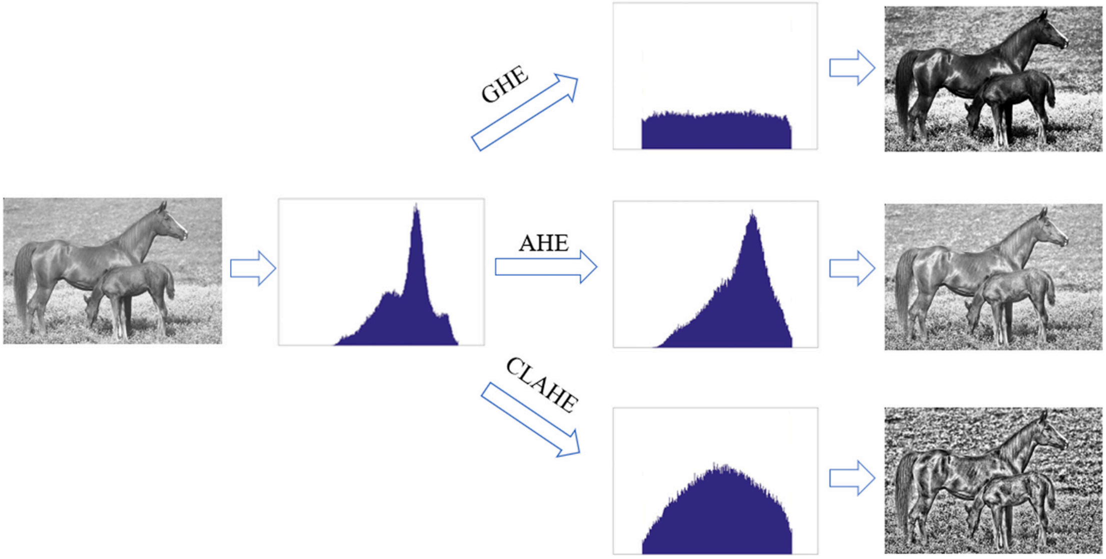



**Fig. 10.** Images processed by the three kinds of HE algorithms.

*2.2.4. Improved HE based methods*

HE algorithm has been widely studied and applied in image enhancement, and there are many algorithms developed based on HE algorithms. Most HE-based algorithms divide the histogram and image into sub-components and perform the histogram equalization operation for the sub-components, respectively. Kim et al. [18] propose a brightness preserving bi-histogram equalization (BBHE), which divides the input image's histogram into two sub histograms according to the mean intensity of the image, and independently equalizes the sub histograms to enhance the image. BBHE can enhance the image contrast while preserving the mean brightness of the input image. Inspired by BBHE, Wang et al. [19] propose an equal-area dualistic sub-image histogram equalization (DSIHE) algorithm, in which the image is decomposed into two equal-area sub-images based on the probability density function, and the two sub-images are equalized respectively. Chen et al. [20] develop a minimum mean brightness error bi-histogram equalization (MMBEBHE) model, which separates the histogram by minimum absolute mean brightness error between the input image and the output image. Based on BBHE, Chen et al. [21] separate each new histogram further based on their respective mean and name the algorithm recursive mean separate histogram equalization (RMSHE). Also, they evaluate the efficiency of the developed algorithm through the scanning electron microscope images. Mohammad et al. [22] propose a segment selective dynamic histogram equalization (SSDHE) which decomposes the histogram of the input image into multiple segments utilizing median values as thresholds. The simulation experiment proves that SSDHE can enhance contrast while preserving the brightness and natural appearance of the images. Kuldeep et al. [23] design an exposure-based HE algorithm (ESIHE). They divide the original image into sub-images of different intensity levels via exposure thresholds, and the individual histogram of sub-images is equalized independently, and finally, all sub-images are integrated into one complete image. Kuldeep et al. [24] introduce a robust contrast enhancement algorithm based on HE algorithm named Median-Mean Based Sub-Image-Clipped Histogram Equalization (MMSICHE). They divide the input image into four sub-images based on the median and mean brightness values and then perform histogram equalization for each sub-image. Similarly, Santhi et al. [25] also divide the input image into four sub-histograms based on its median, and then each partitioned histogram is equalized independently. Tang et al. [26] propose an adaptive image enhancement based on the Bi-Histogram Equalization (AIEBHE) technique, which divides the input histogram into two sub-histograms according to the threshold of the histogram median. In AIEBHE, histogram clipping is performed to control the enhancement rate, and then the clipped sub-histograms are equalized and integrated to obtain the enhanced image.

Most of the improved HE algorithms are designed for the mean brightness of images. Many researchers present use the clipping method to improve the HE algorithm to preserve the mean brightness. Chen et al. [27] propose bi-histogram equalization with a plateau level (BHEPL). BHEPL first divides the input image into two parts based on the mean value, and then two plateau limits are calculated from two sub histograms to avoid over-enhancement. MMSICHE [24] also clips the histogram before dividing the image into sub-images. Tang et al. [28] propose a novel approach based on bi-histogram equalization, named BHEMHB, which segments the input histogram based on the median brightness of the image and alters the histogram bins before HE is applied. Wang et al. [29] design a novel histogram equalization method, CEFPBHE, consisting of adaptive gamma transform, exposure-based histogram splitting, and histogram addition. The object of gamma transform is to restrain histogram spikes from avoiding over-enhancement and noise artifacts effects. Histogram splitting is for preserving mean brightness, and histogram addition is used to control histogram pits. Mun et al. [30] propose a new adaptive plateau limit and a new edge-enhancing transformation function, and a further improved HE algorithm. Upendra et al. [31] propose a novel adaptive image enhancement technique based on genetic algorithm (GAAHE) to enhance magnetic resonance images. Similar work includes [32,33], both of which combine Particle Swarm Optimization (PSO) and HE, and apply them to magnetic resonance images.

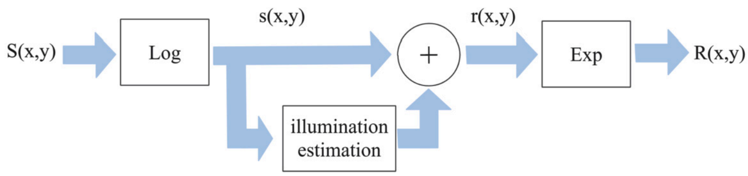



**Fig. 11.** The basic process of Retinex.

### 2.3. Retinex methods

Retinex [34,35], proposed by Land, Edwin H, is a commonly used image enhancement method based on scientific experiments and analysis. The word "Retinex" is a portmanteau formed from "retina" and "cortex". The Retinex model is based on three assumptions:

1. The real world is colorless, and the color is the result of the interaction between light and objects. For example, the water in people's eyes is colorless, but the water-soap film is colorful, resulting from light interference on the surface of the film.
2. Each color region is composed of red, green, and blue primary colors of a given wavelength.
3. The three primary colors determine the color of each unit region.

The theory of Retinex is based on the idea that the color of an object is determined by its ability to reflect long-wave (red), medium-wave (green), and short-wave (blue) illumination rather than the absolute value of the reflected light. Different from the traditional linear and nonlinear methods that can only enhance a certain attribute of images, Retinex can strike a balance in dynamic range compression, edge enhancement, and color invariance, so it can enhance various types of images adaptively. After years of research and development, the Retinex algorithm has been improved from single-scale Retinex algorithm (SSR) [36] to multi-scale Retinex algorithm (MSR) [37], and then to multi-scale Retinex algorithm with color restoration (MSRCR) [38].

#### 2.3.1. Single-scale Retinex (SSR)

According to the theory of Retinex, a given image $S(x,y)$ can be decomposed into two different components: the reflection component $R(x,y)$ and the illumination component $L(x,y)$, and the decomposition can be expressed as $S(x,y) = R(x,y) \cdot L(x,y)$.

The basic idea of the Retinex theory is to remove or reduce the effects of illumination and preserve the essential characteristics of the object. Solving $R(x,y)$ can be regarded as a singular problem, and the basic process is shown in Fig. 11.

The estimation of reflection components by SSR algorithm can be calculated by Eq. (8).

$$\begin{aligned} r(x,y) &= \log S(x,y) - \log[F(x,y) * S(x,y)], \\ r(x,y) &= \log R(x,y), \end{aligned} \tag{8}$$

in which $(x,y)$ is the coordinates of the pixels, $*$ is the convolution operator, $S(x,y)$ and $R(x,y)$ represent the input image and the output image respectively, and $F(x,y)$ represents the Gaussian surround function. $F(x,y)$ can be calculated by Eq. (9).

$$F(x,y) = \lambda e^{\frac{-\left(x^2+y^2\right)}{c^2}}, \tag{9}$$

where $c$ is a Gaussian surround scale which determines the depth of the Retinex scale, and it is usually between 80 and 100. $\lambda$ is a normalization factor given by Eq. (10).

$$\iint F(x,y)dxdy = 1. \tag{10}$$

From the above formulas, the convolution operation in SSR can be viewed as the calculation of the illumination intensity of the image. The physical meaning of SSR can be considered as reducing the illumination of images by calculating the weighted average of the pixels in the image and the surrounding area. The process of the SSR algorithm can be summarized as follows:

1. Read the input image $S(x,y)$. If $S(x,y)$ is a grayscale image, transform the gray value of each pixel of the image to a floating-point number and transform it to the log-domain. If $S(x,y)$ is an RGB image, process each color channel of the image processed separately.
2. Set the parameter $c$, and calculate $\lambda$.
3. According to the above formulas, calculate $r(x,y)$. If the input image is an RGB image, each color channel has a reflection component $r_i(x,y)$.
4. Transform $r(x,y)$ from log-domain to real-domain, and obtain the output image $R(x,y)$.

#### 2.3.2. Multi-scale Retinex (MSR)

Among dynamic range compression and color restoration, the SSR algorithm can only improve one function at the expense of the other one. Jobson et al. [37] propose a multi-scale Retinex algorithm that combines the enhancement results at different scales

(a) Original image (b) SSR (c) MSR (d) MSRCR

**Fig. 12.** Enhanced results by the SSR, MSR, MSRCR.

linearly and takes into account the local and global information. The main idea of MSR is to estimate the illumination component by combining several central surround functions of different scales. The MSR algorithm can be expressed as Eq. (11).

$$r(x,y)=\sum_{k}^{K} w_k\{\log S(x,y)-\log[F(x,y)*S(x,y)]\}, \tag{11}$$

where, $K$ is the number of central surround functions. When $K=1$, MSR is the same as SSR. Normally, $K$ is set to 3, so that the high, medium and low scales can be considered. Variable $w_k$ is the weighting coefficient of the $k$th scale, and it needs to satisfy Eq. (12):

$$\sum_{k=1}^{K} w_k = 1. \tag{12}$$

Benefiting from multi-scale fusion, MSR not only enhances the detail and contrast of the image but also takes color consistency into account.

#### 2.3.3. Multi-scale Retinex algorithm with color restoration (MSRCR)

When enhancing the RGB images by SSR and MSR, the three channels of the image are processed independently, which may lead to the color distortion of the enhanced images. Based on the SSR and MSR algorithm, Jobson et al. [38–40] propose a color restoration multi-scale Retinex algorithm (MSRCR). The expression of MSRCR is as shown in Eq. (13):

$$r_i(x,y)=\sum_{k}^{K} C_i(x,y) w_k\{\log S(x,y)-\log[F(x,y)*S(x,y)]\}. \tag{13}$$

Compared to MSR, the most important improvement of MSRCR is the addition of the color restoration function $C_i(x,y)$, as shown in Eq. (14), and $i$ represents the $i$th channel. Jobson et al. have tried several different color restoration functions for processing on the experimental scene, including linear and nonlinear functions. Through comparative experiments, they found that the following function can provide the best overall color restoration

$$C_i(x,y)=\beta\left(\log\left(\alpha I_i(x,y)\right)-\log\left(\sum_{i\in\{r,g,b\}} I_i(x,y)\right)\right), \tag{14}$$

in which $\beta$ is the gain constant, and $\alpha$ controls the nonlinear degree. MSRCR algorithm applies color restoration factor $C_i$ to adjust the proportion relationship between the three color channels in the original image to highlight the information of relatively dark areas and reduce the defects of image color distortion. Figs. 12(a) to 12(d) show the images enhanced by SSR, MSR, and MSRCR, respectively.

#### 2.3.4. Other improved Retinex based algorithms

The core objective of the Retinex algorithm is to decompose the input image into the reflectance component and the illumination component. Since the introduction of the Retinex theory, multiple methods have been proposed to estimate this illumination effect, and many enhancement algorithms have been proposed based on the Retinex theory. Elad et al. [41] present a non-iterative Retinex algorithm with two special bilateral filters, in which the first evaluates the illumination, and the other is used for the computation of the reflectance. Li et al. [42] propose a new Retinex algorithm based on a recursive bilateral filter, which can effectively deal with the slow processing speed of the bilateral Retinex algorithm. Kimmel et al. [43] propose a variational model for the Retinex problem. In their work, A variational expression is obtained by defining the optimal illumination as the solution of a Quadratic Programming (QP) optimization problem, and they introduce an efficient algorithm based on QP solvers and the fact that the unknown illumination is spatially smooth. In [44], the authors establish a total variation (TV) and nonlocal TV regularized model of Retinex theory that can be solved by a fast computational approach based on Bregman iteration. On the basis of the total variation model, Ng et al. [45] add some constraints and a fidelity term to the Retinex algorithm, which guarantees the existence of the solution for the proposed model. Wang et al. [46] propose a variational model with barriers for Retinex, which is defined as a constrained optimization problem associated with a deduced energy functional by adding two barriers. Fu et al. [47] propose a weighted variational model to estimate both the reflectance and the illumination from an observed image. Their method can preserve the estimated reflectance with more details and suppress noise to some extent. Morel et al. [48] formalize the original Retinex algorithm as a partial differential equation

(a) Original image (b) Enhanced by Retinex

**Fig. 13.** Comparison between the original image and processed by Retinex.

(PDE), and convert the problem into a fast algorithm involving just one parameter. Marcelo et al. [49] provide a new interpretation for the original construction of Retinex. On the basis of their analysis, they present a Kernel-Based Retinex (KBR), which relies on the computation of the expectation value of a suitable random variable weighted with a kernel function. Xuesong et al. [50] propose a Complexity Reduction Retinex model for the enhancement of low luminance retinal fundus images. To improve the computational efficiency, they divide the illumination and reflection components into two independent sub-problems and solve them efficiently by Alternating Direction Minimizing (ADM) method. Guo et al. [51] propose a simple and effective method, named LIME, to enhance low-light images. LIME first estimates the illumination of each pixel individually by finding the maximum value in R, G, and B channels. Then, they propose an augmented Lagrangian multiplier (ALM) based algorithm to exactly solve the refinement problem and design a sped-up solver to intensively reduce the computational load. LIME belongs to the Retinex-based category, and the difference is that LIME only estimates illumination, which shrinks the solution space and reduces the computational cost to reach the desired result. Xutong et al. [52] propose a robust low-light enhancement approach, called LR3M, which is the first to inject low-rank prior into a Retinex decomposition process to suppress noise in the reflectance map.

Yamakawa et al. [56] present an image fusion technique using source image and Retinex-processed image in order to implement high visibility in both bright and dark areas. Jang et al. [57] propose a novel pixel-level multisensor image fusion algorithm with simultaneous contrast enhancement. In order to accomplish both image fusion and contrast enhancement simultaneously, they suggest a modified framework of the subband-decomposed multiscale Retinex (SDMSR). To reduce the color distortion by the dominant chromaticity of the original image, Jang et al. [58] propose a multi-scaled Retinex using a modified local average image. Fu et al. [59] present a new probabilistic method for image enhancement based on simultaneous estimation of illumination and reflectance in the linear domain instead of the logarithmic domain. Petro et al. [60] offer analysis and implementation of Multiscale Retinex and point out and resolve some ambiguities of the method. They also improve the color correction in MSR and propose a multiscale Retinex algorithm with chromaticity preservation (MSRCP). Lin et al. [55] replace the logarithm function in MSR with a customized sigmoid function to minimize data loss, named sigmoid-MSR. The experiment shows that sigmoid-MSR can preserve areas with normal or intensive lighting and suppress noise speckles in extremely low light areas when applied to nighttime images. Furthermore, Matin et al. [61] apply the particle swarm optimization algorithm (PSO) method to adjust the parameters for MSRCP.

The main idea of Retinex is to restore the real image of an object by splitting the input image into the illumination component and reflection component. Retinex methods can not only improve the contrast and brightness of the image but also achieve a balance in dynamic range compression, edge enhancement, and color constancy. However, the images with large brightness differences may contain haloes after being enhanced by Retinex. In addition, the common disadvantages of MSR include lack of edge sharpening, abrupt shadow boundary, distortion of colors, unclear texture, no significant improvement in the details of the highlighted area, and low sensitivity to the highlighted area, etc. To solve these problems, researchers have tried various methods to improve the Retinex algorithm, such as adding bilateral filters [41,42], applying variational models [43–47], using fusion methods [56,57], and so on.

Different traditional methods for image enhancement have different application scenarios. This paper summarizes the purpose and merits of each method, as shown in Table 1.

## 3. Machine learning based methods

Traditional image enhancement methods often bring problems after adjusting the color, brightness, and contrast of the image, such as amplifying noise, the loss of details, and color distortion. For instance, as shown in Fig. 13(a) and Fig. 13(b), although the Retinex algorithm enhances the contrast of the image and makes the objects in the image easily recognizable, the image has a single tone with an overall grayish color. In recent years, as deep learning methods have been successfully applied in many computer-vision tasks, such as face recognition [3] and target detection [4], deep learning has also been widely applied in image enhancement by many scholars. The image enhancement method for low-light images based on deep learning is a kind of data-driven method, which allows the model to automatically learn the features of the images under normal light conditions and reduce the effect on the image caused by low light.

### 3.1. End-to-end learning methods

The end-to-end learning process is a deep learning process in which all parameters are trained jointly rather than step by step. The most common structure of image enhancement algorithms based on deep learning is the encoder-decoder structure, as shown in Fig. 14. LLNet [62] is the first algorithm that enhances low-light images based on deep learning and achieves remarkable results,

**Table 1**
Comparison of traditional algorithms for image enhancement.

| Category | Method | Purpose | Merits |
|---|---|---|---|
| Gray Transformation | Linear Gray Transformation [12] | Adjust the gray value by linear function | |
| | Piece-wise Gray Transformation [12] | Adjust the gray value according to the preset intervals | Suitable for the images with local dark or bright areas |
| | Logarithmic Transformation [53] | Adjust the gray value by logarithmic function | Expands dark pixels of the image and compresses bright pixels |
| | Gamma Transformation [54] | Adjust the gray value by gamma function | Not only suitable for the low-light images, but also suitable for high-light images |
| Histogram Equalization | GHE [14] | Enhance images by balancing gray values | Especially effective for images with concentrated gray values |
| | AHE [14] | Separate an image into several sub-blocks, and process them by histogram equalization, respectively. | Enhance the local contrast and details of the image |
| | CLAHE [17] | Clip the histogram of each sub-block | Limits the amplification of noise |
| | BBHE [18] | Divides the histogram into two sub-histograms according to the mean intensity of the image | Enhance the image contrast while preserving the mean brightness of the input image |
| | DSIHE [19] | Decompose the image into two equal-area sub-images based on the probability density function | Can preserve the mean brightness of the input image better than BBHE |
| | MMBEBHE [20] | Separates the histogram by minimum absolute mean brightness error between the input image and the output image | Suitable for the images with very low, very high and medium mean brightness |
| | RMSHE [21] | Separates the histogram of images recursively based on their respective means | The mean brightness of the output images would meet to the mean brightness of the input images |
| | ESIHE [23] | Divide the original image into sub-images of different intensity levels via exposure thresholds | Show better performance in terms of image visual quality, entropy preservation and contrast enhancement. |
| | AIEBHE [26] | Adaptively selects the smallest value among histogram bins, mean, and median values | outperforms in terms of detail preservation and mean brightness preservation. |
| | BHEPL [27] | Clip the sub-histograms based on the calculated plateau value | Avoid excessive enhancement |
| | CEFPBHE [29] | Merge adaptive gamma transform, exposure-based histogram splitting, and histogram addition | Avoid over-enhancement and noise artifacts effect |
| Retinex | SSR [36] | Decompose an image into two different components: the reflection component and the illumination component | Reduce the effects of illumination and preserve the essential features of the object. |
| | MSR [37] | Estimate the illumination component by combining several different scales central surround function | Balance local and global dynamic range compression |
| | MSRCR [38] | Add a color restoration function | Reduce the defects of image color distortion |
| | Elad et al. [41] | Apply two specially tailored bilateral filters to evaluates the illumination and the reflectance | Effectively handle the edges in the illumination that causes hallow effects |
| | Li et al. [42] | Use recursive bilateral filter to estimate the illumination image | Effectively deal with the slow processing speed of the bilateral Retinex algorithm |
| | Kimmel et al. [44] | Formulate illuminating estimation as a Quadratic Programming optimization problem | Outstand in computational efficiency and parameter robustness |
| | Fu et al. [47] | Estimate both the reflectance and the illumination by a weighted variational model | Preserve the estimated reflectance with more details and suppress noise |
| | Morel et al. [48] | Formalize the original Retinex algorithm as a partial differential equation | Present a fast algorithm involving just one parameter |
| | LIME [51] | Estimate illumination of each pixel individually by finding the maximum value in R, G, and B channels | Outperform several state-of-the-art methods. |
| | MSRCP [55] | Mapp the data to each channel in proportion to the original RGB | Enhance the image while preserving the original color distribution |

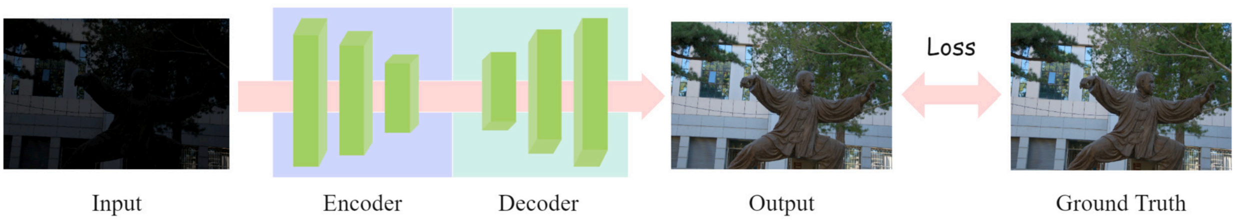



**Fig. 14.** Basic end-to-end learning methods structure.

in which a variant of the stacked-sparse denoising autoencoder is employed to learn from synthetically darkened and noise-added training examples and adaptively enhances images taken from a natural low-light environment or are hardware-degraded. Ren et al. [63] try to enhance the visibility of low-light images based on a trainable hybrid network, where an encoder-decoder network is employed to estimate the global content and a novel spatially variant recurrent neural network (RNN) is employed as an edge stream to model edge details. In [64], the authors also use an encoder-decoder convolution network to build a low-light image enhancement model, and they utilize multi-scale feature maps and join jump connections to avoid gradient disappearance.

Tao et al. [65] propose a CNN-based model to denoise low-light images with a bright channel prior to estimating the transmission parameter. Inspired by MSR [37] and CNN, Shen et al. [66] consider that MSR is equivalent to a feedforward convolutional neural network with different Gaussian convolution kernels and propose an MSR-net that directly learns an end-to-end mapping between dark and bright images. Gharbi et al. [67] focus their research on real-time performance and introduce a new neural network architecture inspired by bilateral grid processing and local affine color transforms. Proved by experiments, their algorithm processes high-resolution images on a smartphone in milliseconds. For underwater image enhancement, Wang et al. [68] propose a CNN-based end-to-end framework UIE-Net, which is trained with two tasks: color correction and haze removal. Besides, the authors synthesize 200000 training images based on the physical underwater imaging model and carry out experiments on benchmark underwater images for cross-scenes. Similarly, Gabriel et al. [69] design a deep CNN to predict HDR values, and Yuen et al. [70] combine Gaussian processing and CNN to enhance the low-light images. Chen et al. [71] develop a pipeline for processing low-light images based on end-to-end training of a fully convolutional network. They compare two network structures: multi-scale context aggregation network (CAN) [72] and U-Net [73], and they finally choose U-net as the default architecture. More importantly, unlike previous work, which was mostly based on synthetic data, they introduce a dataset of raw short-exposure low-light images with corresponding long-exposure reference images. To solve the problem of loss of edge information in dehazing methods, Mingliang et al. [74] propose a gradient-guided dual-branch network for image dehazing in which they explore the hazy image gradient map to guide the model to focus on the hazy regions and edge restoration.

Zhang et al. [75] present a novel attention-based [76] neural network to generate high-quality enhanced low-light images from the raw sensor data. They employ a spatial attention module to focus on denoising by taking advantage of the non-local correlation in the image and a channel attention module to guide the network to refine redundant color features. Atoum et al. [77] propose a color-wise attention network (CWAN), which learns an end-to-end mapping between low-light and enhanced images while searching for any useful color cues in the low-light image to aid in the color enhancement process for low-light image enhancement based on convolutional neural networks. For the low resolution of the captured medical images, Jianrun et al. [78] propose a gated multi-attention feedback network. They introduce a layer attention feature extraction (LAFE) module to refine the feature map and a channel-space attention reconstruction (CSAR) module to enhance the representational ability of the semantic feature map.

### 3.2. Decomposition-based learning methods

Motivated by the excellent model explicable of Retinex theory [34], lots of research work on image enhancement are carried out via combining the idea of image decomposition and deep learning algorithms, such as CNN [66,79–81]. The basic idea of decomposition-based learning methods is divided into two steps. First, the low-light image and normal image are decomposed into reflectance and illumination components via a decomposition module. The components of the low-light image could be optimized by learning from the components of the normal images. Second, the components are further optimized through an adjustment module, and the enhanced image can be obtained by combining the components. Fig. 15 shows the basic structure of decomposition-based learning methods.

Baslamisli et al. [79] analyze the best of deep learning and traditional methods for image processing. First, they propose a physics-based convolutional neural network, IntrinsicNet, which employs the dichromatic reflection mode [82] as a standard reflection model to steer the training process of CNN. Then, they propose the RetiNet, which is a two-stage Retinex-inspired convolutional neural network that first learns to decompose image gradients into intrinsic image gradients, i.e., reflectance and shading gradients. In the second stage, these intrinsic gradients are used to learn the CNN to decompose, at the pixel, the full image into its corresponding reflectance and shading images. LightenNet [80] is a kind of Retinex-based CNN structure that can predict mapping relations between weakly illuminated images and the corresponding illumination map, and the advantage of LightenNet is that it is easy to be trained. Zhang et al. [83] design a network with three subnetworks, called KinD. In KinD, the Layer Decomposition Net is used to decompose images into reflectance components and illumination components, Reflectance Restoration Net is used to restore the reflectance maps of low-light images, and Illumination Adjustment Net is used to flexibly convert one light condition to another separately. Furthermore, the authors propose an improved KinD, named KinD++ [84], which can remove artifacts hidden in images by a multi-scale illumination attention module. Wenhan et al. [85] also propose a deep learning model with multiple subnetworks. In their work, A Sparse Gradient Minimization subnetwork (SGM-Net) is constructed to remove the low-amplitude structures and preserve major edge information. After the learned decomposition, two sub-networks (Enhance-Net and Restore-Net) are utilized to predict the enhanced illumination and reflectance maps, respectively. To fully use scene-level contextual dependencies on spatial scales, Long et al. [86] develop a novel context-sensitive decomposition connection to bridge the reflectance and illumination estimation module. To preserve the color consistency, Zhao et al. [87] decompose an image into a grayscale map and a color histogram. In their model, the grayscale map is used to generate reasonable structures and textures, and the corresponding color histogram is beneficial to keeping color consistency.

Retinex-Net [88] is presented by Wei et al., which consists of a Decom-Net for splitting the input image into lighting-independent reflectance and structure-aware smooth illumination and an Enhance-Net for illumination adjustment. Besides, they build a large-scale dataset with paired low/normal-light images captured in real scenes. To learn an image-to-illumination mapping, Wang et al. [89] present a new neural network for enhancing underexposed photos and design a loss function that adopts constraints and priors on the illumination. Park et al. [90] propose a dual self-encoder network model based on Retinex theory, which combines a stacked

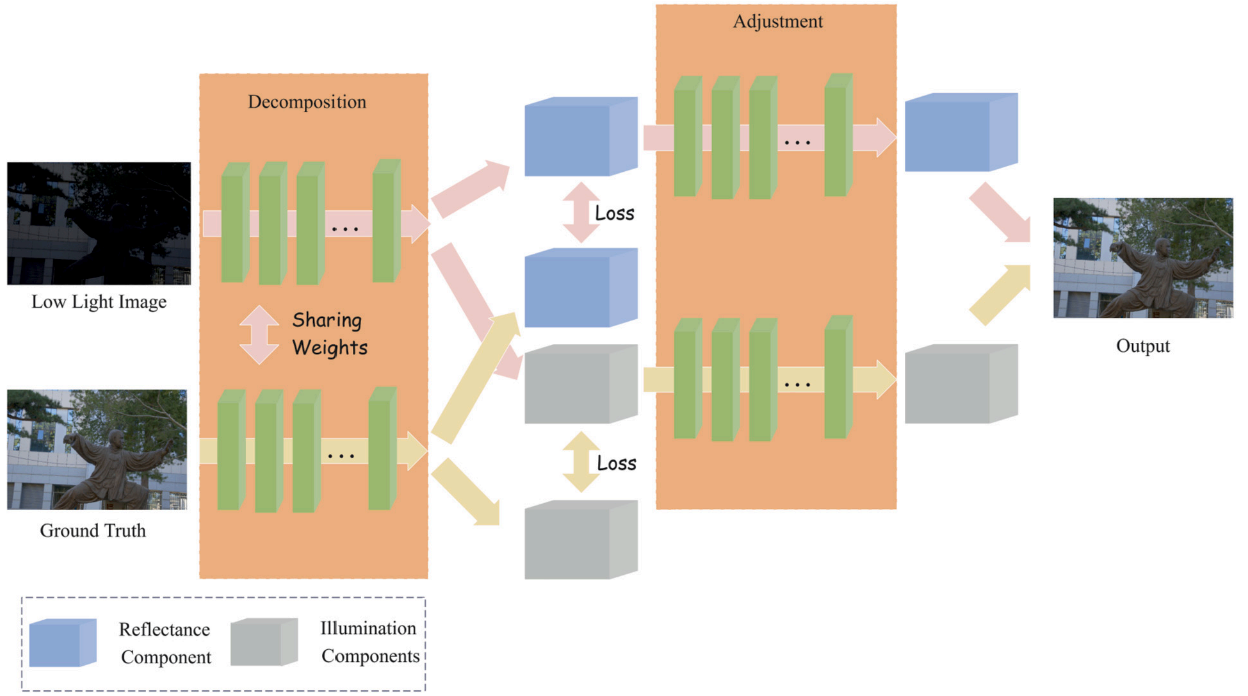



**Fig. 15.** The basic structure of decomposition-based learning methods.

self-encoder with a convolutional self-encoder to realize low-light level enhancement and noise reduction. Wenjing et al. [91] propose a Global illumination-Aware and Detail-preserving Network (GLADNet), which first calculates a global illumination estimation for the low-light input, then adjust the illumination under the guidance of the estimation and supplements the details using a concatenation with the original input. Zhu et al. [92] present a three-branch convolution neural network, namely RRDNet, to decompose the input image into three components: illumination, reflectance, and noise. By formulating the decomposition problem as an implicit prior regularized model, Wenhui et al. [93] propose a Retinex-based deep unfolding network (URetinex-Net). URetinex-Net contains three learning-based modules which are responsible for data-dependent initialization, high-efficient unfolding optimization, and user-specified illumination enhancement, respectively. By assuming that an image can be decomposed into texture and color components, Xiaojie et al. [94] decompose the RGB colorspace into a luminance and a chrominance space. They design an adjustable noise suppression network to eliminate noise in the luminance component and a chrominance mapper to restore colors.

In [95], the authors propose a Retinex Decomposition Network (RDNet) for decomposition, a Fusion Enhancement Network (FENet) for fusion, and a new Generative Adversarial Network (GAN) loss based on Retinex decomposition. Shi et al. [96] also design propose a novel approach for processing low-light images based on the Retinex theory and generative adversarial network. Huang et al. [97] propose a light image enhancement model based on attention mechanism and Retinex. The first estimates the illumination mask of the input image and guides the network to predict the illumination distribution. Then, they apply a module with an attention mechanism to predict the illumination map, and the initial enhanced image is estimated based on the Retinex model. Finally, they modify the color distortion and suppress noise with convolution layers to obtain enhanced results. Fan et al. [98] integrate Retinex theory and the idea of semantic segmentation to construct a pipeline for low-light image enhancement. Liu et al. [99] present a Retinex-inspired unrolling with architecture search (RUAS) to construct a low-light enhancement network. They first establish models to characterize the intrinsic underexposed structure of low-light images based on Retinex theory and unroll optimization processes to construct a holistic propagation structure. Then, a cooperative reference-free learning strategy is designed to discover low-light prior architectures from a compact search space. Furthermore, they design a differentiable strategy to improve RUAS [100], which is able to discover cooperative scene and task architectures from a compact search space. Long et al. [101] build a cascaded illumination learning process with weight sharing to estimate illumination, in which they design a self-calibrated module that realizes the convergence between results of each stage. The method only uses a single basic block for inference, significantly reducing computation costs.

### 3.3. Fusion-based learning methods

Image fusion is the technique of combining multiple images into one that preserves the aspects of relevance of each image [102]. The methods based on image fusion usually take images under different exposure conditions as input or obtain multiscale features by different feature extraction methods. Multiple exposure fusion-based image enhancement normally combines multiple derived images to recover details and resolve color biases, as shown in Fig. 16.

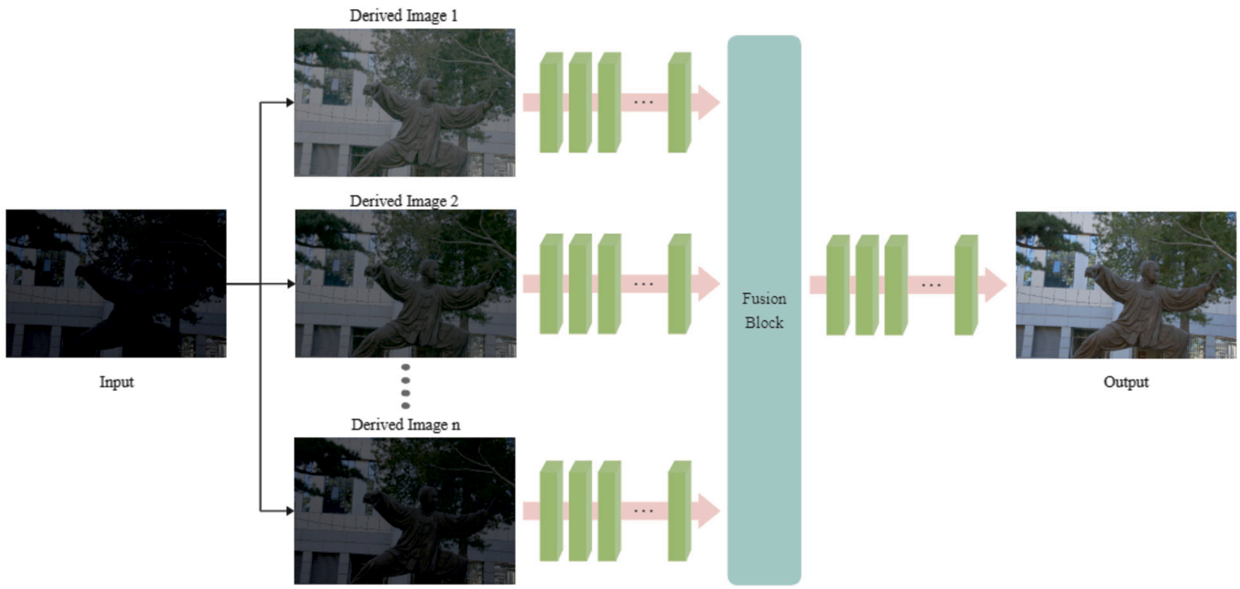



**Fig. 16.** The basic structure of multiple exposure fusion based learning methods.

Jianrui et al. [103] compare and analyze the advantages and disadvantages of single image contrast enhancement (SICE) methods and multi-exposure image fusion (MEF) methods. Based on their analysis, they design a method based on CNN to learn a SICE enhancer, which is able to enhance the low-contrast images with different exposure levels automatically. Lu et al. [104] present a two-branch exposure-fusion network, named TBEFN, by which two enhanced images can be obtained. The final result can be obtained by fusing the two images. Zhu et al. [105] propose a deep learning structure called EEMEFN, which consists of two stages: a novel multi-exposure fusion module together with fusion blocks to combine generated images with different light conditions, and an edge enhancement module to enhance images with sharp edges and fine structures. Inspired by the human visual system, Zhenqiang et al. [106] design a multi-exposure fusion framework for low-light image enhancement, which can adjust the exposure and generate a multi-exposure image set by simulating the human eye. Yu et al. [107] develop a novel algorithm for learning local exposures with deep reinforcement adversarial learning. They first segment an image into sub-images according to dynamic range exposures, and then the local exposure of each sub-image can be learned by reinforcement learning. The final result can be obtained by the fusion of each local best exposure image. Ren et al. [108] develop a fusion-based encoder-decoder network by applying White Balance (WB), Contrast Enhancing (CE), and Gamma Correction (GC). Based on exposure, global contrast, and local contrast, Xu et al. [109] apply a pyramid fusion scheme to fuse the artificial multi-exposure images layer by layer. Cheng et al. [110] propose a deep fusion network (DFN), which is based on CNN, to fuse images created by multiple base image enhancement techniques, including Bright Channel Enhancement, Log Correction, and CLAHE [17]. Lihua et al. [111] propose a symmetric encoder-decoder with residual block (SEDRFuse) network to fuse infrared and visible images for night vision applications. Kui et al. [112] propose a novel degradation-to-refinement generation network, called DRGN. The algorithm first applies a two-step generation network for degradation learning and content refinement, and then constructs a multi-resolution fusion network to represent the target information in a multi-scale collaborative manner.

Models that extract different image features through several branching networks and then fuse them are also classified as fusion-based learning methods. Lv et al. [113] propose a CNN-based method MBLLEN, which consists of three types of modules, i.e., the feature extraction module (FEM), the enhancement module (EM), and the fusion module (FM). They prove that MBLLEN works well in terms of suppressing image noise and artifacts in the low-light regions. Wang et al. [114] design a feature extraction block to extract features and a feature fusion block to fuse multi-level features, and then they apply a channel attention module to estimate the channel importance of input features. Kuang et al. [115] propose an effective nighttime vehicle detection system that combines a novel bioinspired image enhancement approach with a weighted feature fusion technique. Yang et al. [116] present a lightweight, adaptive feature fusion network for image enhancement, consisting of multiple branches with different kernel sizes to generate multi-scale feature maps. Similarly, Liu et al. [117] propose a multi-scale feature fusion-based neural network for image enhancement, which takes into account both global and local features. For the problem that the image enhancement algorithm may overexpose the normal-light areas of the image, Haoyuan et al. [118] propose a local color distribution embedded module to formulate local color distributions in multi-scales to model the correlations across different regions, and a dual-illumination learning mechanism to enhance the regions.

### 3.4. Unpaired learning methods

Most deep learning-based image enhancement algorithms require paired datasets. However, collecting paired images of the same scene in both low and normal light conditions is sometimes difficult, and training a deep learning model based on paired dataset may

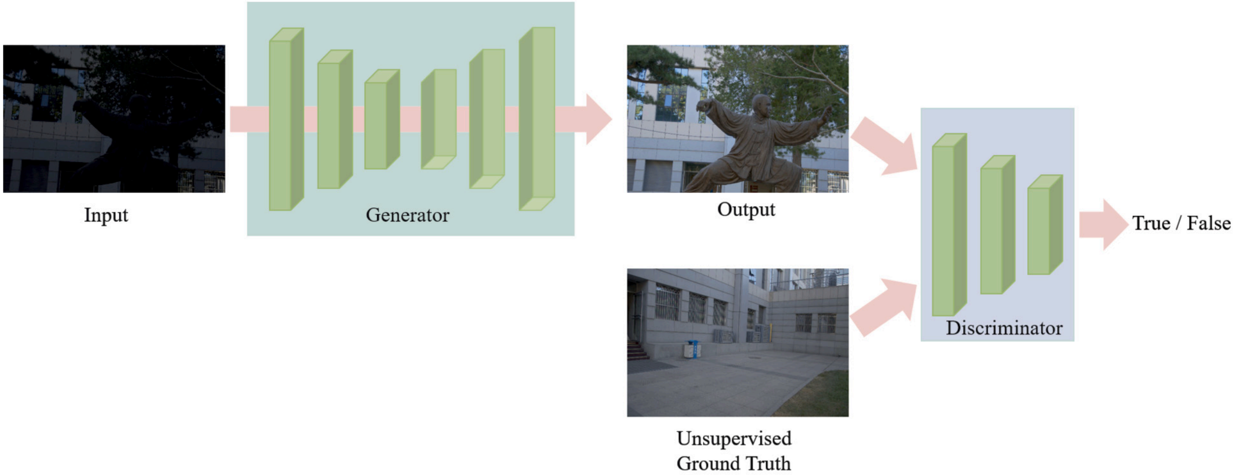



**Fig. 17.** The basic structure of unpaired learning methods.

result in overfitting and limited generalization capability. In contrast to paired learning, some researchers try to adopt unsupervised learning methods to complete image enhancement tasks without paired datasets, as shown in 17. Based on information entropy theory and the Retinex model, Zhang et al. [119] propose a self-supervised low-light image enhancement method that can be trained with low-light images only. EnlightenGAN [120] is an unsupervised learning method based on GAN [121], which can be trained without low/normal-light image pairs. EnlightenGAN adopts an attention-guided U-net [73] as the generator and uses the dual-discriminator to direct the global and local information. Experiments on various datasets demonstrate that EnlightenGAN is easily adaptable to enhancing real-world images from various domains. The framework of EnlightenGAN is shown in Fig. 17. Xiong et al. [122] propose to learn a two-stage GAN-based Framework, including illumination enhancement and noise suppression, to enhance real-world low-light images in a fully unsupervised fashion. Yang et al. [123] design a semi-supervised deep recursive band network to connect fully supervised and unsupervised frameworks. Similarly, researches on GAN-based low-light image enhancement include [124–129].

Some researchers consider image enhancement as a kind of image-specific mapping estimation, and deep learning will be adapted to calculate the best mapping so that the model can be trained without the paired datasets. Guo et al. [130] present a novel method, Zero-Reference Deep Curve Estimation (Zero-DCE), which formulates light enhancement as a task of image-specific curve estimation with a deep network. In order to realize the zero-reference training, they design four non-reference loss functions which implicitly measure the enhancement quality and drive the learning of the network. They further present an accelerated and light version of Zero-DCE, called Zero-DCE++ [131]. Zhang et al. [132] design a small image-specific CNN, namely ExCNet, to estimate the “S-curve [133]” that best fits the test back-lit image. With S-curve, the back-lit image can then be restored accordingly. In [92], the weights of RRDNet will be updated by a zero-shot scheme of iteratively minimizing a specially designed loss function, which is devised to evaluate the current decomposition of the test image and guide noise estimation.

### 3.5. Loss function

The commonly used loss functions include $L_1$, $L_2$, smooth $L_1$, and SSIM loss. Suppose $I$ and $\hat{I}$ are the ground truth image and the predicted image, respectively, and $I_p$ and $\hat{I}_p$ are a pixel of the ground truth image and the predicted image, respectively. The loss functions are defined as Eqs. (15), (16), (17), (18), and (19):

1. $L_1$ loss:

$$L_1 = \sum_{p}^{n} \left| I_P - \hat{I}_p \right| \tag{15}$$

2. $L_2$ loss:

$$L_2 = \sum_{p}^{n} \left( I_P - \hat{I}_p \right)^2 \tag{16}$$

3. smooth $L_1$ loss:

$$smoothL_1 = \frac{1}{n} \sum_{p}^{n} z(I_p, \hat{I}_p), \tag{17}$$

in which

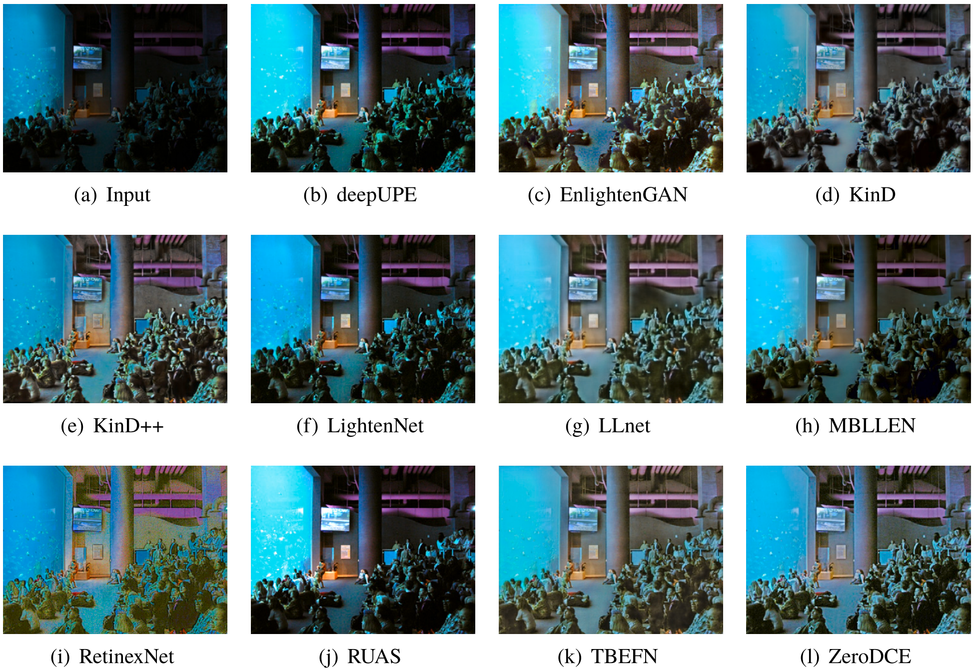



**Fig. 18.** Enhanced results by the machine learning based methods on a sample from [136].

$$z(I_p, \hat{I}_p) = \begin{cases} 0.5\left(I_p - \hat{I}_p\right)^2, & \text{if } \left|I_p - \hat{I}_p\right| < 1 \\ \left|I_p - \hat{I}_p\right| - 0.5, & \text{otherwise} \end{cases}. \tag{18}$$

4. SSIM loss:

$$L_{SSIM} = 1 - SSIM(I, \hat{I}). \tag{19}$$

$L_{SSIM}$ uses the Structural Similarity (SSIM) [134] term which is a commonly-used metric for image reconstruction tasks, which is introduced in detail in Section 4.2.

In addition to these commonly applied loss functions, researchers also introduce some other loss functions according to their model structures and research targets. For example, in [88], the authors introduce an illumination smoothness loss to smooth the textural details and preserve the overall structure boundary. Similar ideas are also adopted in [89,98,130,135]. In [120], to build an unsupervised generative adversarial network, the authors apply the adversarial loss to train their model. In [119], Yu et al. design a loss function based on Retinex theory to train the self-learning model with low-light images only.

### 3.6. Enhanced results by machine learning based methods

In this section, we show the enhanced results of several representative machine learning-based methods. Figs. 18(a) to 18(l) show an indoor multi-person scene and the enhanced images by the methods. Although all the people become clear, these images have problems of overexposure, underexposure, and color distortion. The image enhanced by RetinexNet [88] shows the serious problem of blurring and artifacts, and the same problem exists in the image enhanced by KinD++ [84]. DeepUPE [89], LightenNet [80], and RUAS [99] show an inability to balance exposure. In particular, in the image enhanced by RUAS [99], the background information is almost completely lost, and we cannot even see the man behind it. Although others have improved contrast and clearness while retaining most of the information, they all have varying degrees of color distortion. For example, the color of the image enhanced by ZeroDCE [130] tends to be bluish, and there are halos in the image enhanced by LLNet [62] because of being under-exposed.

Figs. 19(a) to 19(l) show an example of an indoor scene with less light. All methods can improve the brightness of the input image, and the layout and details can be seen clearly. However, none of them successfully recovers the real scene relieved by low-light conditions. In particular, RetinexNet [88] causes obvious color distortion, and ringing artifacts in the image. The images enhanced by EnlightenGAN [120], KinD [83], KinD++ [84], LLNet [62], MBLLEN [113], TBEFN [104], ZeroDCE [130], RUAS [99] tend to over-exposed. The images enhanced by DeepUPE [89], RUAS [99] and LightenNet [80] tend to be under-exposed, in which the details are still hard to distinguish.

Figs. 19(b) to 19(l) and 20(a) show an outdoor example. Similarly, all methods can obviously improve the brightness and contrast of the input image. However, Most algorithms have problems with color distortion and noise generation when dealing with the sky at night in the background. The color distortion and artifacts are still serious in RetinexNet [88] and the KinD++ [84]. The

**Table 2**
Summary of performance for deep learning based image enhancement methods.

| Method | Backbone | Training Dataset | Testing Dataset | PSNR | SSIM | NIQE |
|---|---|---|---|---|---|---|
| LLnet [62] | Deep Autoencoder | Synthetic images | Synthetic images | 19.81 | 0.67 | |
| Ren et al. [63] | RNN | MIT-Adobe FiveK | Synthetic images | 28.43 | 0.96 | |
| LLED-Net [64] | Deep Autoencoder | BSD500 [145] | ImageNet | 27.89 | 0.95 | |
| Tao et al. [65] | CNN | Real images | Real images | | 0.85 | |
| MSR-net [66] | CNN | UCID dataset [146], BSD dataset [145], Google image search | UCID dataset [146], BSD dataset [145], Google image search | | 0.92 | 3.46 |
| Gharbi et al. [67] | CNN | MIT-Adobe FiveK | MIT-Adobe FiveK | 28.40 | | |
| Chen et al. [71] | FCN | SID | SID_Sony camera | 28.88 | 0.79 | |
| | | | SID_Fuji camera | 26.61 | 0.68 | |
| Zhang et al. [75] | Attention-based neural network | Synthetic images | Synthetic images | 20.84 | 0.82 | |
| | | | SID | 27.96 | 0.77 | |
| Atoum et al. [77] | Attention-based neural network | SID | SID_Sony | 28.56 | 0.91 | |
| | | | SID_Fuji | 26.77 | 0.91 | |
| | | | PASCAL_1000 | 29.08 | 0.92 | |
| LightenNet [80] | CNN | Synthetic images | Synthetic images | 21.71 | 0.93 | |
| KinD [83] | Unet | LOL dataset | LOL dataset | 20.87 | 0.80 | |
| KinD + + [84] | Unet | LOL dataset | LOL dataset | 21.30 | 0.82 | |
| | | | MIT-Adobe FiveK | 21.99 | 0.80 | |
| Wenhan et al. [85] | Residual dense network | Synthetic and real images | Synthetic and real images | 22.05 | 0.91 | |
| Retinex-net [88] | Three sub-networks | LOL dataset | LOL dataset | | | |
| DeepUPE [89] | Feature map | MIT-Adobe FiveK | MIT-Adobe FiveK | 30.80 | 0.89 | |
| Park et al. [90] | Dual autoencoder network | Synthetic images | Synthetic images | 17.02 | 0.70 | |
| GLADNet [91] | Encoder-decoder network, CNN | Synthetic images | LIME, DICM, MEF | | | |
| RRDNet [92] | Three-branch CNN | MEF, LIME, DICM, NPE | MEF, LIME, DICM, NPE | | | 3.279 (mean) |
| RDGAN [95] | Two Unet based sub-networks | SICE | SICE | 22.34 | | |
| Yangming et al. [96] | GAN | LOL dataset | LOL dataset | 31.31 | 0.88 | |
| Huang et al. [97] | Attention based network | synthetic images | synthetic images | 23.52 | 0.86 | |
| Fan et al. [98] | Semantic Segmentation network | synthetic dataset | synthetic dataset | 28.82 | 0.95 | 3.05 |
| RUAS [99] | Cooperative Architecture Search | MIT-Adobe FiveK | MIT-Adobe 5K | 20.83 | 0.85 | |
| | | LOL dataset | LOL dataset | 18.23 | 0.72 | |
| RUAS⁺ [100] | Cooperative Architecture Search | MIT-Adobe FiveK, | MIT-Adobe 5K | 21.02 | 0.86 | |
| | | LOL dataset | LOL dataset | 18.2 | 0.72 | |
| URetinex-Net [93] | Two-branch CNN | MIT-Adobe 5K | MIT-Adobe 5K | 21.33 | 0.83 | |
| | | LOL dataset | LOL dataset | 18.95 | 0.78 | |
| Bread [94] | Unet based network | LOL dataset | LOL dataset | 22.96 | 0.84 | 3.95 |
| SCI [101] | Self-calibrated module | MIT-Adobe 5K | MIT-Adobe 5K | 20.45 | 0.89 | 3.96 |
| CSDNet [86] | Unet based network | LOL | LOL | 21.64 | 0.85 | |
| | | | MIT-Adobe 5K | 18.48 | 0.85 | |
| DCC-Net [87] | Three Unet based sub-networks | LOL | LOL | 22.72 | 0.81 | |
| Jianrui et al. [103] | CNN | SCIE | SCIE | 19.77 | 0.93 | |
| TBEFN [104] | Two branch encoder-decoder network | SCIE, LOL | LOL | 17.14 | 0.76 | 3.21 |
| EEMEFN [105] | Two branch Unet based network | SID | SID_Sony camera | 29.60 | 0.796 | |
| | | | SID_Fuji camera | 27.38 | 0.72 | |
| DeepExposure [107] | Reinforced Adversarial Learning | MIT-Adobe FiveK | MIT-Adobe FiveK | 28.38 | | |
| Cheng et al. [110] | CNN | synthetic dataset | synthetic dataset | 24.12 | 0.90 | |
| MBLLEN [113] | Multi-branch Unet based network | synthetic dataset | synthetic dataset | 26.56 | 0.89 | |
| Wang et al. [114] | Unet based network | synthetic dataset | synthetic dataset | 29.68 | | |
| | | | SID | 29.79 | | |
| DRGN [112] | GAN | LOL dataset | LOL dataset | 19.88 | 0.89 | |
| Haoyuan et al. [118] | Two branch Unet based network | MSEC [147] | MSEC [147] | 22.30 | 0.86 | |
| Zhang et al. [119] | Maximum entropy model | LOL dataset | LOL dataset | 19.15 | 0.71 | 4.79 |
| EnlightenGAN [120] | Unet based GAN | images from several datasets | images from several datasets | | | 3.39 |
| Xiong et al. [122] | Two branch GAN | images from [120] | MIT-Adobe FiveK | 19.78 | 0.82 | |
| | | | LOL dataset | 20.04 | 0.82 | |
| DRBN [123] | Recursive network | LOL dataset | LOL dataset | 20.13 | 0.83 | |
| Zero-DCE [130] | Unet based network | SICE | SCIE | 16.57 | | 0.59 |

enhancement performance of RUAS [99] and MBLLEN [113] are not as significant as other algorithms. LightenNet produces obvious halos so that the image becomes blurring. The images enhanced by LLNet [62], ZeroDCE [130] and TBEFN [104] look like they are in a foggy environment since the presence of noise. Table 2 summarizes the frameworks, datasets and performances of the image enhancement algorithms based on deep learning. Because the evaluation methods used in these papers are different, this paper only counts the three most used methods: PSNR, SSIM, and NIQE, which have been introduced in Section 4.

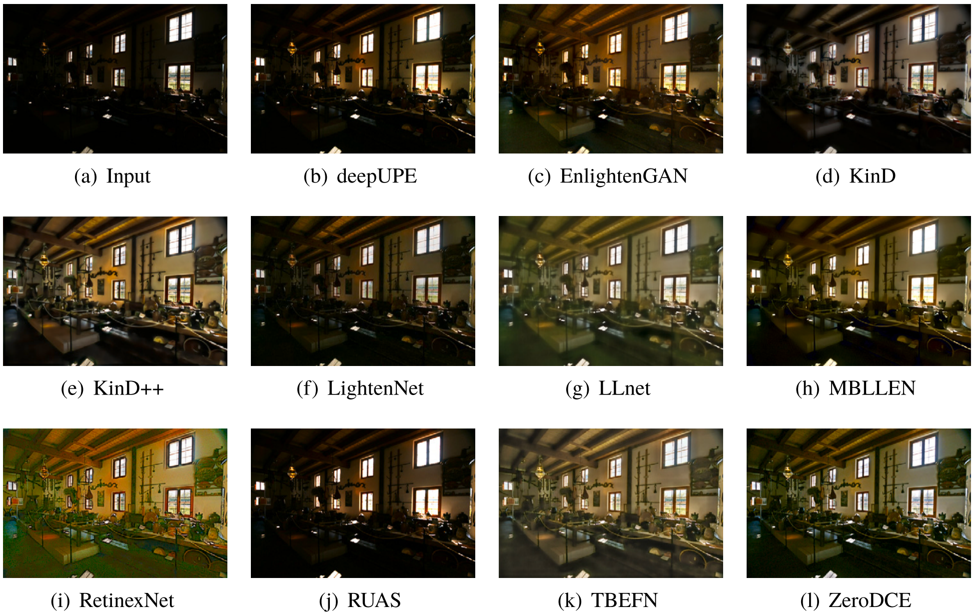



**Fig. 19.** Enhanced results by the machine learning based methods on a sample from [137].

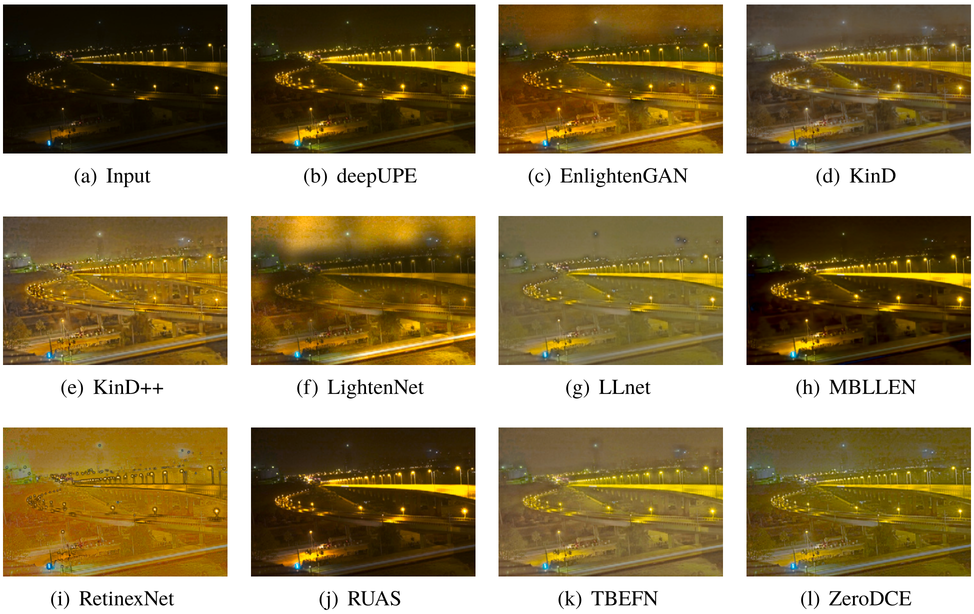



**Fig. 20.** Enhanced results by the machine learning based methods on a sample from [138].

### *3.7. Applications on vision tasks*

In computer vision, image enhancement is often used as the data processing of high-level tasks. In this section, we introduce some work on high-level tasks for low-light images, including object detection and semantic segmentation.

**Object Detection.** Yukihiro et al. [139] propose a domain adaptation method to merge SID [71], and YOLO [140]. The experimental results show that their method can work in scenes illuminated by less than 1 lux. To detect the objects in images under adverse weather conditions, Wenyu et al. [141] propose a novel Image-Adaptive YOLO (IA-YOLO) framework. They design a differentiable image processing module to consider the adverse weather conditions for a YOLO detector based on a small convolutional neural network. Jiaju et al. [142] design a space-time non-local module that leverages the spatial-temporal information across an image sequence in the feature space. They build a robust object detection method for photon-limited conditions through the space-time non-local module and a knowledge distillation module. Xiaojie et al. [143] propose a deep self-adaptive network to detect moving objects in low-light conditions. They design a graph-based unsupervised feature selection module, an anti-occlusion and multi-object handling module, and a weakly fine-tuning strategy. Sobbahi et al. [144] propose a low-light image enhancement model called LL-HFNet (Low-light Homomorphic Filtering Network), which performs image-to-frequency filter learning and is designed for seamless integration into classification models.

**Semantic Segmentation.** Christos et al. [148] design a generic uncertainty-aware annotation and evaluation framework for semantic segmentation in adverse conditions, which explicitly distinguishes invalid from valid regions of input images and applies it to nighttime. Furthermore, they propose a curriculum framework [149] to gradually adapt semantic segmentation models from day to night which exploit cross-time-of-day correspondences between daytime images from a reference map and dark images to guide the label inference in the dark domains. Qi et al. [150] propose a novel Curriculum Domain Adaptation method (CDAda) to realize the smooth semantic knowledge transfer from daytime to nighttime. Xinyi et al. [151] propose a novel domain adaptation network (DANNet) for nighttime semantic segmentation without using labeled nighttime image data. Their model employs adversarial training with a labeled daytime dataset and an unlabeled dataset that contains coarsely aligned day-night image pairs. Huan et al. [152] propose a novel domain adaptation framework via cross-domain correlation distillation (CCDistill). The invariance of illumination or inherent difference between two images could be fully explored to compensate for the lack of labels for low-light images.

## 4. Datasets and evaluation

This section first reviews datasets used in image enhancement, then performance evaluation indices and finally provides an evaluation of image enhancement methods.

### 4.1. Datasets

Many researchers usually share the low illumination image data sets they collect when studying image enhancement algorithms, which provides convenience for other scholars and further research. This paper lists several commonly used image enhancement datasets.

**MIT-Adobe FiveK Dataset** [153]. MIT-Adobe FiveK Dataset contains 5000 images taken with SLR cameras. The dataset covers a broad range of scenes, subjects, and lighting conditions, and each captured image is subsequently retouched by five human artists.

**LOL** [88]. LOL contains 500 low/normal-light image pairs of size 400×600 saved in RGB format, and it is the first dataset containing image pairs taken from real scenes for low-light enhancement.

**SID** [71]. The See-in-the-Dark (SID) dataset contains 5094 raw short-exposure images in both indoor and outdoor environments, and the number of corresponding distinct long-exposure reference images in SID is 424. The outdoor images were generally captured at night, under moonlight or street lighting, in which the illuminance is generally between 0.2 lux and 5 lux. The illuminance at the camera in the indoor scenes is generally between 0.03 lux and 0.3 lux. Images were captured using two cameras: Sony $\alpha$7S II and Fujifilm X-T2. These cameras have different sensors: the Sony camera has a full-frame Bayer sensor, and the Fuji camera has an APS-C X-Trans sensor. The resolution is 4240×2832 for Sony and 6000×4000 for the Fuji images.

**MEF** [137]. Multi-exposure image fusion (MEF) contains 136 fused images, including indoor and outdoor views, natural sceneries, and man-made architectures. All of image sequences contain at least 3 input images that represent underexposed, overexposed, and in-between cases. However, these datasets are in small scale and contain limited scenes.

**SCIE** [103]. SCIE dataset includes 589 sequences from indoor and outdoor scenes, containing a total number of 4,413 multi-exposure images, so that each sequence has 3 to 18 low-contrast images of different exposure levels. Seven types of consumer grade cameras are used to collect the image sequences, including Sony $\alpha$7RII, Sony NEX-5N, Canon EOS-5D Mark II, Canon EOS-750D, Nikon D810, Nikon D7100 and iPhone 6s, and the resolution of most images are between 3000×2000 and 6000×4000.

**NPE** [138]. NPE dataset consists of 46 images captured using the Cannon digital camera, and 110 images downloaded from the websites of some organizations/companies, such as NASA and Google. All the images of the dataset have low contrast in local areas but serious illumination variation in global space.

**ExDARK** [154]. The Exclusively Dark (ExDARK) dataset is a collection of 7,363 low-light images from very low-light environments to twilight (i.e. 10 different conditions) with 12 object classes (similar to PASCAL VOC) annotated on both image class level and local object bounding boxes. It is usually used in the research of object detection in low-light environment.

### 4.2. Image quality assessment

The goal of image quality assessment (IQA) is the construction of computational models that predict the perceived quality of visual images, and it plays an important role in image acquisition, management, communication, and processing systems [155].

**Natural Image Quality Evaluator (NIQE)** [156] NIQE does not require the subjective evaluation score of the original image. It extracts image features from the original images based on a simple spatial domain natural scene statistic (NSS) model, and then fits these features to a multivariate Gaussian model. NIQE is an unsupervised image evaluation method, which does not require the normal image to participate in the calculation. The calculation formula is as shown in Eq. (20).

$$NIQE = \sqrt{\left( (v_1 - v_2)^T \left( \frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (v_1 - v_2) \right)}, \tag{20}$$

in which, $v_1, v_2$ and $\sum_1, \sum_2$ are the mean vectors and covariance matrices of the natural multivariate Gaussian model and the distorted image's multivariate Gaussian model, respectively.

**Mean Square Error (MSE)** MSE represents the direct deviation between the enhanced image and the original image. In an image quality evaluation, a smaller MSE value indicates higher similarity between the enhanced and original images. The formula of MSE shown in Eq. (21).

$$MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \left[ f(i,j) - f_e(i,j) \right]^2, \tag{21}$$

where $f(i,j)$ is the ground-truth normal-light image, and $f_e(i,j)$ is the enhanced image. When calculating the pixel difference at each position, the result is only related to two pixel values at the current position, so it ignores the local structure information of images.

**Peak Signal-to-Noise Ratio (PSNR) [157]** PSNR is commonly used to quantify reconstruction quality for images. The larger the PSNR value is, the smaller the difference between the enhanced images and the reference images. An excessively low PSNR may suggest that the image is distorted. The formula of PSNR is shown in Eq. (22).

$$PSNR = 10 \lg \frac{f_{\max}^2}{MSE}, \tag{22}$$

where $f_{max} = 255$ is the maximum gray value.

**Structural Similarity Index Metric (SSIM) [157]** SSIM is a perception-based method used to measure the similarity of two images, as shown in Eq. (23). It considers image degradation as the perceptional change of structural information and also includes important perceptional phenomena, including luminance masking and contrast masking. The difference with other techniques, such as MSE or PSNR, is that these methods estimate absolute error. Structural similarity means that pixels have a strong interdependence, especially when they are spatially close. These dependencies carry important information about the structure of objects in the visual scene. It measures image similarity from brightness, contrast, and structure, respectively.

$$\begin{aligned} l(x,y) &= \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1}, \\ c(x,y) &= \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2}, \\ s(x,y) &= \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3}, \end{aligned} \tag{23}$$

where $\mu_x$ is the mean value of image $x$, $\mu_y$ is the mean value of image $y$, $\sigma_x^2$ is the variance of image $x$, $\sigma_y^2$ is the variance of image $y$, and $\sigma_{xy}$ is the covariance of $x$ and $y$. $c_1$, $c_2$, and $c_3$ are constant. Normally, in order to avoid 0 in the denominator, $c_1 = \left(k_1 L\right)^2$, $c_2 = \left(k_2 L\right)^2$, $c3 = c2/2$. $L$ is the rang of pixels, and $k_1 = 0.01$, $k_2 = 0.03$ are default values. The SSIM can be expressed as Eq. (24).

$$SSIM = l(x,y) \cdot c(x,y) \cdot s(x,y). \tag{24}$$

**Feature Similarity Index Metric (FSIM) [158]** FSIM, designed by comparing the low-level feature sets between the reference image and the distorted image, combines the features of phase congruency (PC) [159] and gradient magnitude (GM) [160,161] to evaluate images. FSIM assumes that not all pixels in an image are of equal importance. For example, pixels on the edge of an object normally are more important to define the structure of an object than pixels in other background areas. The formula of FSIM is shown in Eq. (25):

$$FSIM = \frac{\sum_{x\in\Omega} S_L(x) \cdot PC_m(x)}{\sum_{x\in\Omega} PC_m(x)}, \tag{25}$$

where Ω means the whole image spatial domain. $PC_m(x)$ is the maximum between the PC of the reference image and the enhanced image. $S_L(x)$ is the metric which defined by combining PC and GM, it can be expressed as Eq. (26).

$$S_L(x) = \left[S_{PC}(x)\right]^\alpha \cdot \left[S_G(x)\right]^\beta, \tag{26}$$

in which $S_PC(x)$ is the PC measure, and $S_G(x)$ is the GM measure. $\alpha$ and $\beta$ are parameters used to adjust the relative importance of PC and GM features, which are set to 1 in [158]. $S_PC(x)$ and $S_G(x)$ are defined as Eq. (27):

$$\begin{aligned} S_{PC}(x) &= \frac{2PC_1(x) \cdot PC_2(x) + T_1}{PC_1^2(x) + PC_2^2(x) + T_1}, \\ S_G(x) &= \frac{2G_1(x) \cdot G_2(x) + T_2}{G_1^2(x) + G_2^2(x) + T_2}, \end{aligned} \tag{27}$$

where $PC_1(x)$ and $PC_2(x)$ are phase congruency of the reference image and the enhanced image respectively, and $G_1(x)$ and $G_2(x)$ are gradient magnitude respectively. $T_1$ is a positive constant to increase the stability of $S_PC(x)$, and $T_2$ depends on the dynamic range of $S_G(x)$.

**Table 3**
Experimental results for traditional methods on LOL dataset.

| Traditional Methods | | NIQE↓ | MSE↑ | PSNR↑ | SSIM↑ | FSIM↑ |
|---|---|---|---|---|---|---|
| Dynamic Range Extension [12] | | 2.997 | 3795 | 14.255 | 0.675 | 0.797 |
| Piece-wise Transformation [12] | | 4.044 | 16438 | 6.600 | 0.006 | 0.566 |
| Inverting [12] | | 3.361 | 19960 | 5.523 | 0.336 | 0.732 |
| Logarithmic transformation [53] | $v = 10$ | 2.809 | 3732 | 12.944 | 0.250 | 0.673 |
| | $v = 30$ | 2.949 | 4349 | 11.938 | 0.0.81 | 0.554 |
| | $v = 100$ | 3.082 | 4248 | 12.108 | 0.078 | 0.574 |
| | $v = 200$ | 3.094 | 3927 | 12.595 | 0.290 | 0.692 |
| Gamma transformation [54] | $\gamma = 0.2$ | 3.287 | 1652 | 17.278 | 0.686 | 0.872 |
| | $\gamma = 0.4$ | 2.975 | 3183 | 15.308 | 0.707 | 0.907 |
| | $\gamma = 2.5$ | 6.010 | 16461 | 6.588 | 0.005 | 0.574 |
| | $\gamma = 5$ | 7.363 | 16597 | 6.549 | 0.003 | 0.554 |
| HE [14] | | 3.848 | 2302 | 15.524 | 0.474 | 0.819 |
| AHE [14] | | 3.243 | 4759 | 11.839 | 0.296 | 0.665 |
| CLAHE [17] | | 3.548 | 8341 | 10.103 | 0.441 | 0.865 |
| SSR [36] | | 3.310 | 2173 | 15.425 | 0.516 | 0.820 |
| MSR [37] | | 4.789 | 2012 | 15.891 | 0.505 | 0.813 |
| MSRCR [38] | | 3.809 | 2162 | 15.789 | 0.667 | 0.887 |

**Table 4**
Experimental results for machine learning based methods on LOL dataset.

| | NIQE↓ | MSE↓ | PSNR↑ | SSIM↑ | FSIM↑ | Parameters (m)↓ | FLOPs (G)↓ | RunTime↓ |
|---|---|---|---|---|---|---|---|---|
| KinD [83] | 3.436 | 1333 | 18.068 | 0.826 | 0.925 | 8.16 | 574.954 | 0.148 |
| KinD+ + [84] | 3.355 | 1201 | 18.232 | 0.808 | 0.871 | 8.275 | 12238.026 | 1.068 |
| DeepUPE [89] | 3.387 | 5008 | 13.160 | 0.541 | 0.901 | 0.567 | - | - |
| EnlightenGAN [120] | 4.467 | 1689 | 18.371 | 0.717 | 0.923 | 8.637 | 273.24 | 0.008 |
| LightenNet [80] | 3.124 | 7432 | 10.534 | 0.442 | 0.858 | 0.028 | - | - |
| LLNet [62] | 4.525 | **1100** | 18.685 | 0.767 | 0.894 | 17.908 | 4124.177 | 36.27 |
| MBLLEN [113] | 3.469 | 1245 | **18.901** | 0.787 | 0.926 | 0.45 | 301.12 | 13.995 |
| RetinexNet [88] | 6.623 | 1254 | 18.062 | 0.539 | 0.864 | 0.555 | 587.47 | 0.12 |
| RUAS [99] | 3.707 | 2279 | 17.129 | 0.701 | 0.893 | 0.003 | 0.281 | 0.006 |
| TBEFN [104] | **2.947** | 1608 | 17.897 | **0.836** | **0.944** | 0.486 | 108.532 | 0.05 |
| ZeroDCE [130] | 3.818 | 7517 | 10.677 | 0.467 | 0.865 | 0.079 | 84.99 | 0.003 |

### *4.3. Evaluation of image enhancement methods*

Table 3 shows the experimental metrics for traditional methods. Since LOL dataset [88] is most commonly used by researchers to evaluate image enhancement algorithms, this paper also applies it to compare and analyze algorithms. Different parameters of the algorithms will generate different images, so only a few parameters are referenced in this paper.

In our experiment, the Piece-wise Transformation makes the low-light area darker and the high-light area, and the experimental images are all low light images, so these images become darker, which can explain why the MSE of Piece-wise Transformation is so large and the SSIM of it is so small. Inverting transforms the high pixel area of the image into a low pixel area and the low pixel area into a high pixel area, so the difference between the processed image and the original image is the largest, which is the reason why the MSE of Inverting are the largest.

The adjustment to the pixel value of an image of Logarithmic transformation is flexible, so its metrics are moderate. When $\gamma > 1$ in Gamma transformation, the pixel values become smaller, so the NIQE and MSE become larger, and PSNR, SSIM, and FSIM become smaller. Unlike the other methods, the aim of HE-based methods is to equalize the pixel distribution of the images instead of just compressing or expanding the pixels of the images. Similarly, the idea of Decomposition-based methods is to make the generated images more suitable for human aesthetics through image decomposition, so the metrics based on pixel difference are normally difficult to reflect the characteristics of these algorithms.

Table 4 shows the experimental metrics for machine learning based methods, including the computational complexity. MSE of LightenNet and ZeroDCE is the largest, and SSIM of it is the smallest among the algorithms, and both of them have relatively few models. It means the generated images enhanced by them are the most different from the normal images. Specifically, LightenNet simply merges Retinex and CNN to construct the image enhancement model without considering color consistency and local information. It also can be obtained that there are artifacts in the enhanced images by LightenNet. The main idea of ZeroDCE is to estimate the deep curve of an image without any paired or unpaired data, which may cause insufficient enhancement. The NIQE of RetinexNet is the largest, which means the images enhanced by it are the most distorted among the algorithms. Subjectively, we can get the same result from the experimental results that the images enhanced by RetinexNet have the greatest color distortion. It can be concluded that image enhancement algorithms inspired by Retinex need to optimize the performance of models in color consistency and local information of images. The FSIM of all algorithms is relatively high, which means these images have a high degree of similarity to the normal image features. The number of parameters in RUAS is minimal, and the results in Section 3 show that the contrast improvement of images enhanced by RUAS is relatively low.

## 5. Conclusion

The purpose of image enhancement algorithm for low-light images is to enlarge the difference between the features of different objects in the image, suppress the features that are not interesting, improve the image quality and enrich the information. This paper discusses the idea, purpose, implementation steps, advantages and shortages of image enhancement from the views of traditional methods and machine learning-based algorithms in detail. Aiming at analyzing the image enhancement algorithms based on machine learning from the digital image theory, we innovatively classify them according to the model strategy and the traditional methods combined with the algorithm. In order to compare different algorithms, we reproduce some algorithms and quantitatively analyze their performance in image quality evaluation methods.

Image enhancement for low-light images plays an important role in computer vision. According to the discussion above, this paper proposes the following potential research directions for image enhancement algorithms. (1) **Unsupervised learning:** In real scenes, there is usually less standard paired data to help researchers train models. Although there are some unsupervised algorithms, they either still have a requirement for data, or have an insufficient performance; (2) **Generalization ability:** In addition to low-light, images often suffer a variety of common corruptions, such as noise, high-contrast, fog, and so on. (3) **High-level application:** Currently, in order to evaluate image enhancement algorithms, researchers usually evaluate their improvement for high-level applications such as object detection. However, more complex applications should be considered, such as depth estimation and 3D reconstruction.

## CRediT authorship contribution statement

All authors listed have significantly contributed to the development and the writing of this article.

## Declaration of competing interest

The authors declare the following conflict of interests: This work is partially supported by the XJTLU AI University Research Centre and Jiangsu (Provincial) Data Science and Cognitive Computational Engineering Research Centre at the Suzhou Science and Technology Project-Key Industrial Technology Innovation (Grant No. SYG202006, SYG202122), Future Network Scientific Research Fund Project (FNSRFP-2021-YB-41), the Key Program Special Fund of Xi'an Jiaotong-Liverpool University (XJTLU), Suzhou, China (Grant No. KSF-A-19, KSF-E-65, KSF-P-02, KSF-E-54).

## Data availability

Data included in article/supplementary material/referenced in article.

## Acknowledgements

Jieming Ma was supported by XJTLU AI University Research Centre and Jiangsu (Provincial) Data Science and Cognitive Computational Engineering Research Centre at the Suzhou Science and Technology Project-Key Industrial Technology Innovation [SYG202006, SYG202122].

## References

[1] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. Van Der Laak, Bram Van Ginneken, Clara I. Sánchez, A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88.

[2] Claudine Badue, Rânik Guidolini, Raphael Vivacqua Carneiro, Pedro Azevedo, Vinicius B. Cardoso, Avelino Forechi, Luan Jesus, Rodrigo Berriel, Thiago M. Paixao, Filipe Mutz, et al., Self-driving cars: a survey, Expert Syst. Appl. 165 (2021) 113816.

[3] Iacopo Masi, Yue Wu, Tal Hassner, Prem Natarajan, Deep face recognition: a survey, in: 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), IEEE, 2018, pp. 471–478.

[4] Zhong-Qiu Zhao, Peng Zheng, Shou-tao Xu, Xindong Wu, Object detection with deep learning: a review, IEEE Trans. Neural Netw. Learn. Syst. 30 (11) (2019) 3212–3232.

[5] Jiahang Liu, Chenghu Zhou, Peng Chen, Chaomeng Kang, An efficient contrast enhancement method for remote sensing images, IEEE Geosci. Remote Sens. Lett. 14 (10) (2017) 1715–1719.

[6] Tong Liu, Zhaowei Chen, Yi Yang, Zehao Wu, Haowei Li, Lane detection in low-light conditions using an efficient data enhancement: light conditions style transfer, in: 2020 IEEE Intelligent Vehicles Symposium (IV), IEEE, 2020, pp. 1394–1399.

[7] Wencheng Wang, Xiaojin Wu, Xiaohui Yuan, Zairui Gao, An experiment-based review of low-light image enhancement methods, IEEE Access 8 (2020) 87884–87917.

[8] Muhammad Tahir Rasheed, Guiyu Guo, Daming Shi, Hufsa Khan, Xiaochun Cheng, An empirical study on Retinex methods for low-light image enhancement, Remote Sens. 14 (18) (2022) 4608.

[9] Chongyi Li, Chunle Guo, Linghao Han, Jun Jiang, Ming-Ming Cheng, Jinwei Gu, Chen Change Loy, Low-light image and video enhancement using deep learning: a survey, IEEE Trans. Pattern Anal. Mach. Intell. 44 (12) (2021) 9396–9416.

[10] Rayan Al Sobbahi, Joe Tekli, Comparing deep learning models for low-light natural scene image enhancement and their impact on object detection and classification: overview, empirical evaluation, and challenges, in: Signal Processing: Image Communication, 2022, p. 116848.

[11] David C.C. Wang, Anthony H. Vagnucci, Ching-Chung Li, Digital image enhancement: a survey, Comput. Vis. Graph. Image Process. 24 (3) (1983) 363–381.
[12] A. Raji, A. Thaibaoui, E. Petit, P. Bunel, G. Mimoun, A gray-level transformation-based method for image enhancement, Pattern Recognit. Lett. 19 (13) (1998) 1207–1212.
[13] V. Rajamani, P. Babu, S. Jaiganesh, A review of various global contrast enhancement techniques for still images using histogram modification framework, Int. J. Eng. Trends Technol. 4 (4) (2013) 1045–1048.
[14] D. Vijayalakshmi, Malaya Kumar Nath, Om Prakash Acharya, A comprehensive survey on image contrast enhancement techniques in spatial domain, Sens. Imaging 21 (1) (2020) 1–40.
[15] Manpreet Kaur, Jasdeep Kaur, Jappreet Kaur, Survey of contrast enhancement techniques based on histogram equalization, Int. J. Adv. Comput. Sci. Appl. 2 (7) (2011).
[16] Hui Zhu, Francis H.Y. Chan, Francis K. Lam, Image contrast enhancement by constrained local histogram equalization, Comput. Vis. Image Underst. 73 (2) (1999) 281–290.
[17] K. Zuiderveld, Contrast limited adaptive histogram equalization, Graph. Gems (1994) 474–485.
[18] Yeong-Taeg Kim, Contrast enhancement using brightness preserving bi-histogram equalization, IEEE Trans. Consum. Electron. 43 (1) (1997) 1–8.
[19] Yu Wang, Qian Chen, Baeomin Zhang, Image enhancement based on equal area dualistic sub-image histogram equalization method, IEEE Trans. Consum. Electron. 45 (1) (1999) 68–75.
[20] Soong-Der Chen, Abd Rahman Ramli, Minimum mean brightness error bi-histogram equalization in contrast enhancement, IEEE Trans. Consum. Electron. 49 (4) (2003) 1310–1319.
[21] Soong-Der Chen, Abd Rahman Ramli, Contrast enhancement using recursive mean-separate histogram equalization for scalable brightness preservation, IEEE Trans. Consum. Electron. 49 (4) (2003) 1301–1309.
[22] Mohammad Farhan Khan, Ekram Khan, Z.A. Abbasi, Segment selective dynamic histogram equalization for brightness preserving contrast enhancement of images, Optik 125 (3) (2014) 1385–1389.
[23] Kuldeep Singh, Rajiv Kapoor, Image enhancement using exposure based sub image histogram equalization, Pattern Recognit. Lett. 36 (2014) 10–14.
[24] Kuldeep Singh, Rajiv Kapoor, Image enhancement via median-mean based sub-image-clipped histogram equalization, Optik 125 (17) (2014) 4646–4651.
[25] K. Santhi, R.S.D. Wahida Banu, Adaptive contrast enhancement using modified histogram equalization, Optik, Int. J. Light Electron Opt. 126 (19) (2015) 1809–1814.
[26] Jing Rui Tang, Nor Ashidi Mat Isa, Adaptive image enhancement based on bi-histogram equalization with a clipping limit, Comput. Electr. Eng. 40 (8) (2014) 86–103.
[27] Chen Hee Ooi, Nicholas Sia Pik Kong, Haidi Ibrahim, Bi-histogram equalization with a plateau limit for digital image enhancement, IEEE Trans. Consum. Electron. 55 (4) (2009) 2072–2080.
[28] Jing Rui Tang, Nor Ashidi Mat Isa, Bi-histogram equalization using modified histogram bins, Appl. Soft Comput. 55 (2017) 31–43.
[29] Xuewen Wang, Lixia Chen, Contrast enhancement using feature-preserving bi-histogram equalization, Signal Image Video Process. 12 (4) (2018) 685–692.
[30] Junwon Mun, Yuneseok Jang, Yoojun Nam, Jaeseok Kim, Edge-enhancing bi-histogram equalisation using guided image filter, J. Vis. Commun. Image Represent. 58 (2019) 688–700.
[31] Upendra Kumar Acharya, Sandeep Kumar, Genetic algorithm based adaptive histogram equalization (GAAHE) technique for medical image enhancement, Optik 230 (2021) 166273.
[32] Boyina Subrahmanyeswara Rao, Dynamic histogram equalization for contrast enhancement for digital images, Appl. Soft Comput. 89 (2020) 106114.
[33] Upendra Kumar Acharya, Sandeep Kumar, Particle swarm optimized texture based histogram equalization (PSOTHE) for MRI brain image enhancement, Optik 224 (2020) 165760.
[34] Edwin H. Land, The Retinex, Am. Sci. 52 (2) (1964) 247–264.
[35] E.H. Land, Lightness and Retinex theory, J. Opt. Soc. Am. 61 (1) (1971) 1–11.
[36] D.J. Jobson, Z.U. Rahman, G.A. Woodell, Properties and performance of a center/surround Retinex, IEEE Trans. Image Process. 6 (3) (1997) 451–462.
[37] Z. Rahman, D.J. Jobson, G.A. Woodell, Multi-scale Retinex for color image enhancement, in: Proceedings of 3rd IEEE International Conference on Image Processing, vol. 3, 1996, pp. 1003–1006.
[38] D.J. Jobson, Z. Rahman, G.A. Woodell, A multiscale Retinex for bridging the gap between color images and the human observation of scenes, IEEE Trans. Image Process. 6 (7) (2002) 965–976.
[39] Zia-ur Rahman, Daniel J. Jobson, Glenn A. Woodell, Retinex processing for automatic image enhancement, J. Electron. Imaging 13 (1) (2004) 100–110.
[40] Zia-ur Rahman, Glenn A. Woodell, Daniel J. Jobson, et al., A comparison of the multiscale Retinex with other image enhancement techniques, in: Is and T Annual Conference, Citeseer, 1997, pp. 426–431.
[41] Michael Elad, Retinex by two bilateral filters, in: International Conference on Scale-Space Theories in Computer Vision, Springer, 2005, pp. 217–229.
[42] Di Li, Yadi Zhang, Pengcheng Wen, Linting Bai, A Retinex algorithm for image enhancement based on recursive bilateral filtering, in: 2015 11th International Conference on Computational Intelligence and Security (CIS), IEEE, 2015, pp. 154–157.
[43] Wenye Ma, Stanley Osher, A TV Bregman iterative model of Retinex theory, Inverse Probl. Imaging 6 (4) (2012) 697.
[44] Ron Kimmel, Michael Elad, Doron Shaked, Renato Keshet, Irwin Sobel, A variational framework for Retinex, Int. J. Comput. Vis. 52 (1) (2003) 7–23.
[45] Michael K. Ng, Wei Wang, A total variation model for Retinex, SIAM J. Imaging Sci. 4 (1) (2011) 345–365.
[46] Wang Wei, Chuanjiang He, A variational model with barrier functionals for Retinex, SIAM J. Imaging Sci. 8 (3) (2015) 1955–1980.
[47] Xueyang Fu, Delu Zeng, Yue Huang, Xiao-Ping Zhang, Xinghao Ding, A weighted variational model for simultaneous reflectance and illumination estimation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2782–2790.
[48] Jean Michel Morel, Ana Belén Petro, Catalina Sbert, A PDE formalization of Retinex theory, IEEE Trans. Image Process. 19 (11) (2010) 2825–2837.
[49] Marcelo Bertalmío, Vicent Caselles, Edoardo Provenzi, Issues about Retinex theory and contrast enhancement, Int. J. Comput. Vis. 83 (1) (2009) 101–119.
[50] Xuesong Li, Mingliang Gao, Jianrun Shang, Jinfeng Pan, Qilei Li, A complexity reduction based Retinex model for low luminance retinal fundus image enhancement, Netw. Model. Anal. Health Inform. Bioinform. 11 (1) (2022) 1–10.
[51] Xiaojie Guo, Yu Li, Haibin Ling, LIME: low-light image enhancement via illumination map estimation, IEEE Trans. Image Process. 26 (2) (2016) 982–993.
[52] Xutong Ren, Wenhan Yang, Wen-Huang Cheng, Jiaying Liu, LR3M: robust low-light enhancement via low-rank regularized Retinex model, IEEE Trans. Image Process. 29 (2020) 5862–5876.
[53] Michel Jourlin, Josselin Breugnot, Frédéric Itthirad, Mohamed Bouabdellah, Brigitte Closs, Logarithmic Image Processing for Color Images, Advances in Imaging and Electron Physics, vol. 168, Elsevier, 2011, pp. 65–107.
[54] Gursharn Singh, Anand Mittal, Various image enhancement techniques-a critical review, Int. J. Innov. Sci. Res. 10 (2) (2014) 267–274.
[55] Haoning Lin, Zhenwei Shi, Multi-scale Retinex improvement for nighttime image enhancement, Optik 125 (24) (2014) 7143–7148.
[56] Masaya Yamakawa, Yasunori Sugita, Image enhancement using Retinex and image fusion techniques, Electron. Commun. Jpn. 101 (8) (2018) 52–63.
[57] Jae Ho Jang, Yoonsung Bae, Jong Beom Ra, Contrast-enhanced fusion of multisensor images using subband-decomposed multiscale Retinex, IEEE Trans. Image Process. 21 (8) (2012) 3479–3490.
[58] In-Su Jang, Kee-Hyon Park, Yeong-Ho Ha, Color correction by estimation of dominant chromaticity in multi-scaled Retinex, J. Imaging Sci. Technol. 53 (5) (2009) 50502–50511.
[59] Xueyang Fu, Yinghao Liao, Delu Zeng, Yue Huang, Xiao-Ping Zhang, Xinghao Ding, A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation, IEEE Trans. Image Process. 24 (12) (2015) 4965–4977.

[60] Ana Belén Petro, Catalina Sbert, Jean-Michel Morel, Multiscale Retinex, in: Image Processing on Line, 2014, pp. 71–88.
[61] Farzin Matin, Yoosoo Jeong, Hanhoon Park, Retinex-based image enhancement with particle swarm optimization and multi-objective function, IEICE Trans. Inf. Syst. 103 (12) (2020) 2721–2724.
[62] Kin Gwn Lore, Adedotun Akintayo, Soumik Sarkar, LLNET: a deep autoencoder approach to natural low-light image enhancement, Pattern Recognit. 61 (2017) 650–662.
[63] Wenqi Ren, Sifei Liu, Lin Ma, Qianqian Xu, Xiangyu Xu, Xiaochun Cao, Junping Du, Ming-Hsuan Yang, Low-light image enhancement via a deep hybrid network, IEEE Trans. Image Process. 28 (9) (2019) 4364–4375.
[64] Qiming Li, Haishen Wu, Lu Xu, Likai Wang, Yueqi Lv, Xinjie Kang, Low-light image enhancement based on deep symmetric encoder–decoder convolutional networks, Symmetry 12 (3) (2020) 446.
[65] Li Tao, Chuang Zhu, Jiawen Song, Tao Lu, Huizhu Jia, Xiaodong Xie, Low-light image enhancement using CNN and bright channel prior, in: 2017 IEEE International Conference on Image Processing (ICIP), IEEE, 2017, pp. 3215–3219.
[66] Liang Shen, Zihan Yue, Fan Feng, Quan Chen, Shihao Liu, Jie Ma, MSR-net: low-light image enhancement using deep convolutional network, arXiv preprint, arXiv:1711.02488, 2017.
[67] Michaël Gharbi, Jiawen Chen, Jonathan T. Barron, Samuel W. Hasinoff, Frédo Durand, Deep bilateral learning for real-time image enhancement, ACM Trans. Graph. 36 (4) (2017) 1–12.
[68] Yang Wang, Jing Zhang, Yang Cao, Zengfu Wang, A deep CNN method for underwater image enhancement, in: 2017 IEEE International Conference on Image Processing (ICIP), IEEE, 2017, pp. 1382–1386.
[69] Gabriel Eilertsen, Joel Kronander, Gyorgy Denes, Rafał K. Mantiuk, Jonas Unger, HDR image reconstruction from a single exposure using deep CNNs, ACM Trans. Graph. 36 (6) (2017) 1–15.
[70] Yuen Peng Loh, Xuefeng Liang, Chee Seng Chan, Low-light image enhancement using Gaussian process for features retrieval, Signal Process. Image Commun. 74 (2019) 175–190.
[71] Chen Chen, Qifeng Chen, Jia Xu, Vladlen Koltun, Learning to see in the dark, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 3291–3300.
[72] Qifeng Chen, Jia Xu, Vladlen Koltun, Fast image processing with fully-convolutional networks, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2497–2506.
[73] Olaf Ronneberger, Philipp Fischer, Thomas Brox, U-Net: convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.
[74] Mingliang Gao, Qingyu Mao, Qilei Li, Xiangyu Guo, Gwanggil Jeon, Lina Liu, Gradient guided dual-branch network for image dehazing, J. Circuits Syst. Comput. (2022).
[75] Cheng Zhang, Qingsen Yan, Yu Zhu, Xianjun Li, Jinqiu Sun, Yanning Zhang, Attention-based network for low-light image enhancement, in: 2020 IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2020, pp. 1–6.
[76] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, Illia Polosukhin, Attention is all you need, in: Advances in Neural Information Processing Systems, 2017, pp. 5998–6008.
[77] Yousef Atoum, Mao Ye, Liu Ren, Ying Tai, Xiaoming Liu, Color-wise attention network for low-light image enhancement, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 506–507.
[78] Jianrun Shang, Xue Zhang, Guisheng Zhang, Wenhao Song, Jinyong Chen, Qilei Li, Mingliang Gao, Gated multi-attention feedback network for medical image super-resolution, Electronics 11 (21) (2022) 3554.
[79] Anil S. Baslamisli, Hoang-An Le, Theo Gevers, CNN based learning using reflection and Retinex models for intrinsic image decomposition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 6674–6683.
[80] Chongyi Li, Jichang Guo, Fatih Porikli, Yanwei Pang, LightenNet: a convolutional neural network for weakly illuminated image enhancement, Pattern Recognit. Lett. 104 (2018) 15–22.
[81] Liyun Zhuang, Yepeng Guan, Image enhancement by deep learning network based on derived image and Retinex, in: 2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), IEEE, 2019, pp. 1670–1673.
[82] Steven A. Shafer, Using color to separate reflection components, Color Res. Appl. 10 (4) (1985) 210–218.
[83] Yonghua Zhang, Jiawan Zhang, Xiaojie Guo, Kindling the darkness: a practical low-light image enhancer, in: Proceedings of the 27th ACM International Conference on Multimedia, 2019, pp. 1632–1640.
[84] Yonghua Zhang, Xiaojie Guo, Jiayi Ma, Wei Liu, Jiawan Zhang, Beyond brightening low-light images, Int. J. Comput. Vis. 129 (4) (2021) 1013–1037.
[85] Wenhan Yang, Wenjing Wang, Haofeng Huang, Shiqi Wang, Jiaying Liu, Sparse gradient regularized deep Retinex network for robust low-light image enhancement, IEEE Trans. Image Process. 30 (2021) 2072–2086.
[86] Long Ma, Risheng Liu, Jiaao Zhang, Xin Fan, Zhongxuan Luo, Learning deep context-sensitive decomposition for low-light image enhancement, IEEE Trans. Neural Netw. Learn. Syst. (2021).
[87] Zhao Zhang, Huan Zheng, Richang Hong, Mingliang Xu, Shuicheng Yan, Meng Wang, Deep color consistent network for low-light image enhancement, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 1899–1908.
[88] Chen Wei, Wenjing Wang, Wenhan Yang, Jiaying Liu, Deep Retinex decomposition for low-light enhancement, arXiv preprint, arXiv:1808.04560, 2018.
[89] Ruixing Wang, Qing Zhang, Chi-Wing Fu, Xiaoyong Shen, Wei-Shi Zheng, Jiaya Jia, Underexposed photo enhancement using deep illumination estimation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 6849–6857.
[90] Seonhee Park, Soohwan Yu, Minseo Kim, Kwanwoo Park, Joonki Paik, Dual autoencoder network for Retinex-based low-light image enhancement, IEEE Access 6 (2018) 22084–22093.
[91] Wenjing Wang, Chen Wei, Wenhan Yang, Jiaying Liu, GLADNet: Low-light enhancement network with global awareness, in: 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), IEEE, 2018, pp. 751–755.
[92] Anqi Zhu, Lin Zhang, Ying Shen, Yong Ma, Shengjie Zhao, Yicong Zhou, Zero-shot restoration of underexposed images via robust Retinex decomposition, in: 2020 IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2020, pp. 1–6.
[93] Wenhui Wu, Jian Weng, Pingping Zhang, Xu Wang, Wenhan Yang, Jianmin Jiang, URetinex-Net: Retinex-based deep unfolding network for low-light image enhancement, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 5901–5910.
[94] Xiaojie Guo, Qiming Hu, Low-light image enhancement via breaking down the darkness, Int. J. Comput. Vis. 131 (1) (2023) 48–66.
[95] Junyi Wang, Weimin Tan, Xuejing Niu, Bo Yan, RDGAN: Retinex decomposition based adversarial learning for low-light enhancement, in: 2019 IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2019, pp. 1186–1191.
[96] Yangming Shi, Xiaopo Wu, Ming Zhu, Low-light image enhancement algorithm based on Retinex and generative adversarial network, arXiv preprint, arXiv:1906.06027, 2019.
[97] Wei Huang, Yifeng Zhu, Rui Huang, Low light image enhancement network with attention mechanism and Retinex model, IEEE Access 8 (2020) 74306–74314.
[98] Minhao Fan, Wenjing Wang, Wenhan Yang, Jiaying Liu, Integrating semantic segmentation and Retinex model for low-light image enhancement, in: Proceedings of the 28th ACM International Conference on Multimedia, 2020, pp. 2317–2325.
[99] Risheng Liu, Long Ma, Jiaao Zhang, Xin Fan, Zhongxuan Luo, Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 10561–10570.

[100] Risheng Liu, Long Ma, Tengyu Ma, Xin Fan, Zhongxuan Luo, Learning with nested scene modeling and cooperative architecture search for low-light vision, IEEE Trans. Pattern Anal. Mach. Intell. (2022).
[101] Long Ma, Tengyu Ma, Risheng Liu, Xin Fan, Zhongxuan Luo, Toward fast, flexible, and robust low-light image enhancement, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 5637–5646.
[102] Mohammad Bagher Akbari Haghighat, Ali Aghagolzadeh, Hadi Seyedarabi, A non-reference image fusion metric based on mutual information of image features, Comput. Electr. Eng. 37 (5) (2011) 744–756.
[103] Jianrui Cai, Shuhang Gu, Lei Zhang, Learning a deep single image contrast enhancer from multi-exposure images, IEEE Trans. Image Process. 27 (4) (2018) 2049–2062.
[104] Lu Kun, Lihong Zhang, TBEFN: a two-branch exposure-fusion network for low-light image enhancement, IEEE Trans. Multimed. (2020).
[105] Minfeng Zhu, Pingbo Pan, Wei Chen, Yi Yang, EEMEFN: low-light image enhancement via edge-enhanced multi-exposure fusion network, Proc. AAAI Conf. Artif. Intell. 34 (2020) 13106–13113.
[106] Zhenqiang Ying, Ge Li, Wen Gao, A bio-inspired multi-exposure fusion framework for low-light image enhancement, arXiv preprint, arXiv:1711.00591, 2017.
[107] Runsheng Yu, Wenyu Liu, Yasen Zhang, Zhi Qu, Deli Zhao, Bo Zhang, DeepExposure: learning to expose photos with asynchronously reinforced adversarial learning, in: Proceedings of the 32nd International Conference on Neural Information Processing Systems, 2018, pp. 2153–2163.
[108] Wenqi Ren, Lin Ma, Jiawei Zhang, Jinshan Pan, Xiaochun Cao, Wei Liu, Ming-Hsuan Yang, Gated fusion network for single image dehazing, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 3253–3261.
[109] Yadong Xu, Cheng Yang, Beibei Sun, Xiaoan Yan, Minglong Chen, A novel multi-scale fusion framework for detail-preserving low-light image enhancement, Inf. Sci. 548 (2021) 378–397.
[110] Yu Cheng, Jia Yan, Zhou Wang, Enhancement of weakly illuminated images by deep fusion networks, in: 2019 IEEE International Conference on Image Processing (ICIP), IEEE, 2019, pp. 924–928.
[111] Lihua Jian, Xiaomin Yang, Zheng Liu, Gwanggil Jeon, Mingliang Gao, David Chisholm, SEDRFuse: a symmetric encoder–decoder with residual block network for infrared and visible image fusion, IEEE Trans. Instrum. Meas. 70 (2020) 1–15.
[112] Kui Jiang, Zhongyuan Wang, Zheng Wang, Chen Chen, Peng Yi, Tao Lu, Chia-Wen Lin, Degrade is upgrade: learning degradation for low-light image enhancement, Proc. AAAI Conf. Artif. Intell. 36 (2022) 1078–1086.
[113] Feifan Lv, Feng Lu, Jianhua Wu, Chongsoon Lim, MBLLEN: low-light image/video enhancement using CNNs, in: BMVC, 2018, p. 220.
[114] Lei Wang, Guangtao Fu, Zhuqing Jiang, Guodong Ju, Aidong Men, Low-light image enhancement with attention and multi-level feature fusion, in: 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), IEEE, 2019, pp. 276–281.
[115] Hulin Kuang, Xianshi Zhang, Yong-Jie Li, Leanne Lai Hang Chan, Hong Yan, Nighttime vehicle detection based on bio-inspired image enhancement and weighted score-level feature fusion, IEEE Trans. Intell. Transp. Syst. 18 (4) (2016) 927–936.
[116] Hao-Hsiang Yang, Kuan-Chih Huang, Wei-Ting Chen, LAFFNet: a lightweight adaptive feature fusion network for underwater image enhancement, arXiv preprint, arXiv:2105.01299, 2021.
[117] Xiaodong Liu, Zhi Gao, M. Ben Chen, MLFFNet: multi-level feature fusion net for underwater image enhancement, in: OCEANS 2019 MTS/IEEE SEATTLE, IEEE, 2019, pp. 1–6.
[118] Haoyuan Wang, Ke Xu, Rynson W.H. Lau, Local color distributions prior for image enhancement, in: European Conference on Computer Vision, Springer, 2022, pp. 343–359.
[119] Yu Zhang, Xiaoguang Di, Bin Zhang, Chunhui Wang, Self-supervised image enhancement network: training with low light images only, arXiv preprint, arXiv:2002.11300, 2020.
[120] Yifan Jiang, Xinyu Gong, Ding Liu, Yu Cheng, Chen Fang, Xiaohui Shen, Jianchao Yang, Pan Zhou, Zhangyang Wang, EnlightenGAN: deep light enhancement without paired supervision, IEEE Trans. Image Process. 30 (2021) 2340–2349.
[121] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, Generative adversarial nets, Adv. Neural Inf. Process. Syst. 27 (2014).
[122] Wei Xiong, Ding Liu, Xiaohui Shen, Chen Fang, Jiebo Luo, Unsupervised real-world low-light image enhancement with decoupled networks, arXiv preprint, arXiv:2005.02818, 2020.
[123] Wenhan Yang, Shiqi Wang, Yuming Fang, Yue Wang, Jiaying Liu, From fidelity to perceptual quality: a semi-supervised approach for low-light image enhancement, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 3063–3072.
[124] Hanyu Li, Jingjing Li, Wei Wang, A fusion adversarial underwater image enhancement network with a public test dataset, arXiv preprint, arXiv:1906.06819, 2019.
[125] Guisik Kim, Dokyeong Kwon, Junseok Kwon, Low-lightgan: low-light enhancement via advanced generative adversarial network with task-driven training, in: 2019 IEEE International Conference on Image Processing (ICIP), IEEE, 2019, pp. 2811–2815.
[126] Yingying Meng, Deqiang Kong, Zhenfeng Zhu, Yao Zhao, From night to day: GANs based low quality image enhancement, Neural Process. Lett. 50 (1) (2019) 799–814.
[127] Yijun Liu, Zhengning Wang, Yi Zeng, Hao Zeng, Deming Zhao, PD-GAN: perceptual-details GAN for extremely noisy low light image enhancement, in: ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2021, pp. 1840–1844.
[128] Ning Rao, Tao Lu, Qiang Zhou, Yanduo Zhang, Zhongyuan Wang, Seeing in the dark by component-GAN, IEEE Signal Process. Lett. (2021).
[129] Jing Wang, Ping Li, Jianhua Deng, Yongzhao Du, Jiafu Zhuang, Peidong Liang, Peizhong Liu, CA-GAN: class-condition attention GAN for underwater image enhancement, IEEE Access 8 (2020) 130719–130728.
[130] 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: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 1780–1789.
[131] Chongyi Li, Chunle Guo, Chen Change Loy, Learning to enhance low-light image via zero-reference deep curve estimation, arXiv preprint, arXiv:2103.00860, 2021.
[132] Lin Zhang, Lijun Zhang, Xiao Liu, Ying Shen, Shaoming Zhang, Shengjie Zhao, Zero-shot restoration of back-lit images using deep internal learning, in: Proceedings of the 27th ACM International Conference on Multimedia, 2019, pp. 1623–1631.
[133] Lu Yuan, Jian Sun, Automatic exposure correction of consumer photographs, in: European Conference on Computer Vision, Springer, 2012, pp. 771–785.
[134] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, Eero P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[135] Feifan Lv, Bo Liu, Feng Lu, Fast enhancement for non-uniform illumination images using light-weight CNNs, in: Proceedings of the 28th ACM International Conference on Multimedia, 2020, pp. 1450–1458.
[136] Chulwoo Lee, Chul Lee, Chang-Su Kim, Contrast enhancement based on layered difference representation of 2D histograms, IEEE Trans. Image Process. 22 (12) (2013) 5372–5384.
[137] Kede Ma, Kai Zeng, Zhou Wang, Perceptual quality assessment for multi-exposure image fusion, IEEE Trans. Image Process. 24 (11) (2015) 3345–3356.
[138] Shuhang Wang, Jin Zheng, Hai-Miao Hu, Bo Li, Naturalness preserved enhancement algorithm for non-uniform illumination images, IEEE Trans. Image Process. 22 (9) (2013) 3538–3548.
[139] Yukihiro Sasagawa, Hajime Nagahara, YOLO in the dark-domain adaptation method for merging multiple models, in: European Conference on Computer Vision, Springer, 2020, pp. 345–359.

[140] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, You only look once: unified, real-time object detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 779–788.

[141] Wenyu Liu, Gaofeng Ren, Runsheng Yu, Shi Guo, Jianke Zhu, Lei Zhang, Image-adaptive YOLO for object detection in adverse weather conditions, Proc. AAAI Conf. Artif. Intell. 36 (2022) 1792–1800.

[142] Chengxi Li, Xiangyu Qu, Abhiram Gnanasambandam, Omar A. Elgendy, Jiaju Ma, Stanley H. Chan, Photon-limited object detection using non-local feature matching and knowledge distillation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 3976–3987.

[143] Xiaojie Huang, Moving object detection in low-luminance images, Vis. Comput. (2021) 1–13.

[144] Rayan Al Sobbahi, Joe Tekli, Low-light homomorphic filtering network for integrating image enhancement and classification, Signal Process. Image Commun. 100 (2022) 116527.

[145] David Martin, Charless Fowlkes, Doron Tal, Jitendra Malik, A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics, in: Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, vol. 2, IEEE, 2001, pp. 416–423.

[146] Gerald Schaefer, Michal Stich, UCID: An Uncompressed Color Image Database, Storage and Retrieval Methods and Applications for Multimedia 2004, vol. 5307, International Society for Optics and Photonics, 2003, pp. 472–480.

[147] Mahmoud Afifi, Konstantinos G. Derpanis, Bjorn Ommer, Michael S. Brown, Learning multi-scale photo exposure correction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 9157–9167.

[148] Christos Sakaridis, Dengxin Dai, Luc Van Gool, Guided curriculum model adaptation and uncertainty-aware evaluation for semantic nighttime image segmentation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 7374–7383.

[149] Christos Sakaridis, Dengxin Dai, Luc Van Gool, Map-guided curriculum domain adaptation and uncertainty-aware evaluation for semantic nighttime image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 44 (6) (2022) 3139–3153.

[150] Qi Xu, Yinan Ma, Jing Wu, Chengnian Long, Xiaolin Huang, CDAda: a curriculum domain adaptation for nighttime semantic segmentation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 2962–2971.

[151] Xinyi Wu, Zhenyao Wu, Hao Guo, Lili Ju, Song Wang, DANNet: a one-stage domain adaptation network for unsupervised nighttime semantic segmentation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 15769–15778.

[152] Huan Gao, Jichang Guo, Guoli Wang, Qian Zhang, Cross-domain correlation distillation for unsupervised domain adaptation in nighttime semantic segmentation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 9913–9923.

[153] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, Frédo Durand, Learning photographic global tonal adjustment with a database of input/output image pairs, in: The Twenty-Fourth IEEE Conference on Computer Vision and Pattern Recognition, 2011.

[154] Yuen Peng Loh, Chee Seng Chan, Getting to know low-light images with the exclusively dark dataset, Comput. Vis. Image Underst. 178 (2019) 30–42.

[155] Zhou Wang, Alan C. Bovik, Modern image quality assessment, Synth. Lect. Image Video Multimed. Process. 2 (1) (2006) 1–156.

[156] Anish Mittal, Rajiv Soundararajan, Alan C. Bovik, Making a “completely blind” image quality analyzer, IEEE Signal Process. Lett. 20 (3) (2012) 209–212.

[157] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, Eero P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.

[158] Lin Zhang, Lei Zhang, Xuanqin Mou, David Zhang, FSIM: a feature similarity index for image quality assessment, IEEE Trans. Image Process. 20 (8) (2011) 2378–2386.

[159] Peter Kovesi, et al., Image features from phase congruency, Videre: J. Comput. Vis. Res. 1 (3) (1999) 1–26.

[160] Ramesh Jain, Rangachar Kasturi, Brian G. Schunck, Machine Vision, vol. 5, McGraw-Hill, New York, 1995.

[161] Bernd Jähne, Horst Haussecker, Peter Geissler, Handbook of Computer Vision and Applications, vol. 2, Citeseer, 1999.