SPECIAL ISSUE PAPER

Low-light-level image enhancement algorithm based on integrated networks

Peng Wang[1,2,3] · Jiao Wu[3] · Haiyan Wang[2,4] · Xiaoyan Li[3] · Yongxia Yang[3]

Abstract

In dark or poorly lit environments, it is often difficult for the naked eye to distinguish low-light-level images because of low brightness, low contrast and noise, and it is difficult to perform subsequent image processing (such as video surveillance and target detection). To solve these problems, this paper proposes a low-light-level image enhancement algorithm based on deep learning. First, the low-light-level image is segmented into several super-pixels, and the noise level of each super-pixel is estimated by the ratio of the local standard deviation to the local gradient. Then, the image is inverted and smoothed by a BM3D filter, and the structural filter adaptive method is used to obtain complete images without noise but with the correct texture. Finally, the noise-free image and texture-complete images are applied to the integrated network, which can not only enhance the contrast but also effectively prevent the over-enhancement of the contrast. The experimental results show that this method is superior to traditional methods in terms of both subjective and objective evaluation, and the peak signal–noise ratio and improved structural similarity are 31.64 dB and 91.2%, respectively.

1 Introduction

Image enhancement is an important branch of image processing. Its purpose is to convert images or videos captured in different environments into clear, high-quality images or videos using effective methods [1]. Because of low contrast, low brightness and varying noise, an image taken at night or in a low-light-level environment has a low dynamic range, which leads to a significant decline in image quality,

difficulty in distinguishing details with the naked eye, and difficulty in carrying out subsequent image processing (such as target detection and video surveillance) [2]. Therefore, low-light-level image enhancement has always been a popular issue in the field of computer vision, and it has increasingly important theoretical significance and research application value [3].

To improve image contrast, brightness and quality, scholars have studied this problem from different angles, such as the histogram equalization (HE) method among traditional methods [4], the method based on retinex theory [5], and the method based on the defogging model [6]. In recent years, with the development of deep learning, remarkable progress has been made in the application of computer vision in many fields [7]; especially in super-resolution, image de-noising, target detection and tracking, deep learning has made breakthroughs, triggering a research boom in academia and industry. Deep learning extracts the features of images or data by building a network model similar to the human brain information-processing mechanism, using efficient learning strategies, fitting complex nonlinear functions and obtaining the expected results. For low-light-level image processing, a convolutional neural network model based on deep learning

✉ Peng Wang
  wang_peng@xatu.edu.cn

1 School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, Shaanxi, China

2 Key Laboratory of Ocean Acoustics and Sensing (Northwestern Polytechnical University), Ministry of Industry and Information Technology, Xi'an 710072, Shaanxi, China

3 School of Electronics and Information Engineering, Xi'an Technological University, Xi'an 710021, China

4 School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an 710021, Shaanxi, China

is proposed for image enhancement [8]. The advantages of three structures, namely, the local receptive field, weight sharing and pooling, not only reduce the number of training parameters and the training difficulty but also have good robustness to distortion-invariant operations on images such as zooming, rotation and translation, which makes the extracted deep-seated features have a stronger generalization ability. Therefore, a low-light all-in-one network (LLAON) is proposed to enhance low-light-level images. Different filters and convolution kernels are used to filter low-light-level images, and then enhanced images are obtained. These enhanced images retain the original features and textures. In addition, the LLAON integrates the mean-square error (MSE) function to reconstruct accurate image textures and improve image quality.

2 Related work

At present, there are four main branches in the field of low-light-level image enhancement: (1) the histogram equalization method, which requires that the pixel values maintain their relative relationships while obeying a uniform distribution, and the adaptive histogram equalization method of limiting contrast, which controls the enhancement range of the local contrast by restricting the height of the local histogram. The local contrast is limited by over-enhancement. This method guarantees a fast processing speed and can effectively improve contrast, but it is prone to color bias, and gray-level merging leads to the loss of detailed information. (2) Based on the retinex theory, this kind of method holds that the image observed by human eyes can be expressed as the product of an illumination component and a reflection component. The illumination component reflects the illumination condition, and the reflection component is an intrinsic attribute of the image. By estimating and removing the illumination component from the original image, the reflection component can be obtained, and image enhancement can be achieved. Therefore, the core step of the retinex algorithm is to accurately estimate the illumination component.

Researchers have adopted different methods to estimate the illumination component and have proposed different retinex algorithms. For example, the single-scale retinex (SSR) algorithm [9], multiscale retinex (MSR) algorithm [10] and MSR algorithm with color restoration (MSRCR) [11]. The SSR algorithm estimates the image using Gauss filtering. The MSR algorithm is a linear weighted sum of multiple scales of SSR. It is better at improving the contrast and brightness of the image than SSR, but it will also make the image edge sharpening inadequate, and some colors will be distorted. The MSRCR algorithm introduces color restoration based on the MSR algorithm so that the enhanced image will not exhibit distortion, but the color of the image

will deviate from the original color, and the overall color will be white. Fu et al. proposed a weighted variational model, namely, SRAIE (simultaneous reflectance and illumination estimation), which estimates both the illumination and reflection components [12]. K Ganga Bhavani proposed the LIME (low-light image enhancement via illumination map estimation) algorithm, which first uses the original low-light-level image red (R), green (G), and blue (B) channels. In the three channels, the maximum value of the illumination image is obtained, and then the original illumination image is continuously revised through the structure priors to obtain the final illumination image [13]. (3) Based on the method of the defogging model, Dong found that the inverted low-light-level image has a high similarity to the fog image and proposed the method of defogging to enhance the low-light-level image, achieving a good enhancement effect. Although this method can improve the visual quality to a certain extent, the enhanced images are often inconsistent with the actual scene and tend to have artifacts at the edges. In addition, there are some algorithms that transform the color space of an RGB (red, green, blue) image to other color spaces to achieve image enhancement, which can effectively maintain the color of the image and prevent serious color distortion of the enhanced image, but the brightness and contrast still need to be improved [14]. (4) With the development of deep learning in image processing, detection and tracking [15], a deep neural network stacked sparse denoising self-encoder is proposed to enhance gray images, which indicates the feasibility of deep learning methods. The convolutional neural network (CNN) is used to enhance low-illumination images, which improves the image contrast and image quality [16].

To solve the above problems, an effective low-light-level image enhancement algorithm based on deep learning is proposed in this paper. The method consists of two stages: an adaptive denoising method based on super-pixel segmentation and all-in-one network contrast enhancement. The super-pixel method can segment a whole image into several super-pixels quickly using the geometric flow method. In the first stage of the algorithm, according to this excellent algorithm, the low-light-level image is divided into many super-pixels, and then the noise texture level of each super-pixel is estimated according to the ratio of the local standard deviation to the local gradient. Then, the image is inverted. The super-pixel segmentation algorithm is used to extract the contour information of noisy images. According to the noise texture characteristics of the super-pixels, the original image is divided into texture changes and locally similar blocks. Then, the smooth bottom layer is extracted using the BM3D filter [17], which has self-adaptability. Finally, the image is inverted. The first-order differential extracts another detail layer and smooths it with a structural filter [18]. The two layers are combined adaptively to obtain the complete

image without noise but with texture. In the second stage, the denoised image is applied to the integrated defogging network; we call this the low-light-level integrated network (LLAON), and it solves the problem of enhancing contrast and preventing excessive contrast enhancement. Finally, the image generated in the steps above is inverted to obtain the enhanced image. The experimental results show that compared with other mainstream algorithms, the proposed algorithm can not only significantly enhance brightness and contrast but also keep the image color information unchanged and further improve the visual perception and objective evaluation indicators.

The main contributions of this work are as follows:

1. A unified framework is proposed to de-noise and enhance contrast jointly to eliminate noise before contrast enhancement. A local adaptive denoising scheme based on super-pixels is proposed, which can eliminate noise while preserving texture details.
2. To overcome the problem of over-enhancement and under-enhancement of contrast, the LLAON is used to enhance contrast. The mean-square error (MSE) function [19] is integrated into the LLAON to reconstruct more accurate image textures and improve image quality. Experiments show that not only the peak signal–noise ratio (PSNR) [20] is improved, but also structural similarity (SSIM) [21] and visual quality are improved.
3. The experimental results show that the proposed method is superior to traditional methods in terms of both subjective and objective evaluation, and the PSNR and SSIM are 31.64 dB and 91.2%, respectively.

3 Low-light-level image enhancement method

In this section, we introduce the concept of the LLAON; Sect. 3.1 gives the overall network architecture; Sect. 3.2 describes the adaptive de-noising method based on super-pixels; Sect. 3.3 describes the integrated network

architecture; and Sect. 3.4 gives the image quality evaluation PSNR and SSIM.

3.1 Overall network architecture

In this paper, an effective low-light-level image enhancement algorithm based on an integrated network is proposed. The overall model architecture is shown in Fig. 1.

In this paper, the low-light-level image enhancement algorithm based on the integrated network is divided into two stages: adaptive de-noising based on super-pixels and contrast enhancement based on an integrated network of low-light-level images. In the first stage of the algorithm, the low-light-level image is divided into many super-pixels, and then the noise texture level of each super-pixel is estimated according to the ratio of the local standard deviation to the local gradient. Then, the image is inverted. The super-pixel segmentation algorithm is used to extract the contour information of noisy images. According to the noise texture characteristics of super-pixels, the original image is divided into texture changes and local similar blocks. Then, the smooth bottom layer is extracted by the BM3D filter, which has self-adaptability. Finally, the image is inverted. The first-order differential extracts another detail layer and smooths it with structural filters. The two layers are combined adaptively to obtain a complete image without noise but with texture. In the second stage, noise-free texture images are applied to the integrated network to form a low-light-level integrated network (LLAON). This method not only effectively enhances the contrast but also prevents the problem of over-enhancement of the contrast. Finally, we inverted the generated image to obtain the final enhanced image.

3.2 Super-pixel self-adaptive de-noising

Due to the characteristics of the human visual system itself, the visibility of noise in different regions of an image varies greatly. For example, compared with the noise in a smooth-textured region, noise is more obvious with a complex texture. On the other hand, low-pass de-noising filters often degrade texture details. Therefore, the intensity of the
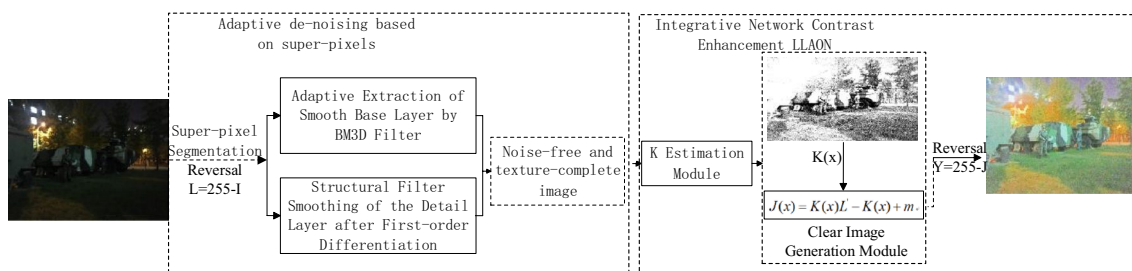

Fig. 1 The overall model of the low-light-level image enhancement algorithm based on an integrated network

adaptive de-noising filter needs to be determined according to the local features of the image region. Strong filters are more suitable for small areas with high-visibility noise or low-complexity details, while weak filters are preferred for small areas with low-visibility noise or high-complexity details. In this paper, a local adaptive denoising scheme based on super-pixels is proposed, which can keep the texture details and eliminate the noise at the same time so that the local noise can be handled well and the feature information will not be lost due to excessive denoising.

First, the low-light-level image I is segmented into several sub-regions using the super-pixel method. It should be noted that the super-pixel segmentation method in this paper adopts the method proposed in [22], SEEDS (super pixels extracted via energy-driven sampling), because it has better retention characteristics for the boundary of the object. For each sub-region, the following method is used to determine the smoothness, assuming that the noise is additive white Gaussian noise (AWGN). The local gradient of the super-pixel is represented by the standard deviation of the super-pixel. After adding AWGN to the clear image, the flat area increases significantly. However, there is not much change in the texture area. For normalized images in the range [0, 1], the standard deviation varies by an order of magnitude. Therefore, taking into account the normalized ratio of the sums to measure the noise texture grade of the quantum region, the following is shown:

The standard deviation of super-pixel i is represented by δ_i, and the local gradient of the super-pixel is represented by ∇_i. After adding AWGN to the clear image, the flat area ∇_i increases significantly. However, there is not much change in texture area ∇_i for normalized images in the range [0, 1], and the standard deviation δ_i varies by an order of magnitude. Therefore, the normalized ratio λ_i between δ_i and ∇_i is considered to measure the noise texture grade of the region. The specific formulas are as follows:

$$\lambda_i = \frac{\delta_i}{\nabla_i}. \tag{1}$$

To facilitate de-noising and enhance the contrast using an integrated network de-noising algorithm, the input image $L = 255 - I$ is inverted by using I. Inspired by the non-sharpening mask filter, the de-noising operation L is defined as L'. L' is obtained from the weighting of the base level and the noise-free level of L.

$$L' = \lambda \cdot m(L) + n(L). \tag{2}$$

Among the variables, $m(L)$ and $n(L)$ represent the noise-free layer and base layer of L, respectively. For sub-regions with small λ values, some details are added to restrict the noise level. For sub-regions with larger λ values, more details are added to the underlying layer. Using a BM3D

filter to smooth the image can obtain the image base level, which leads to the desirable result that the AWGN can be effectively reduced. The noise texture horizontal coefficient λ is used as the weight to generate the roots.

$$m(L) = \lambda \cdot m^{\text{fine}}(L) + (1 - \lambda) \cdot m^{\text{coarse}}(L). \tag{3}$$

In the formula, $m^{\text{fine}}(L)$ and $m^{\text{coarse}}(L)$ represent the smoothing results of BM3D filters using parameters that are half and twice the local standard deviation δ_i of the average value of the low-light-level image I.

To obtain the detail layer $n_1(L)$, we select the first-order differential of the inverted image L. It is found that the random noise in detail layer $n_1(L)$ tends to fuse the image detail texture. Therefore, we use structured filters to smooth the texture detail layer of the image to retain useful image texture to obtain smoother and more complete texture feature results $n(L)$.

3.3 Integrative network image enhancement

Because the low-light image L' is similar to a defogged image, an effective integrated network is used to enhance the contrast. One of the definitions of the atmospheric scattering model is:

$$L' = J(x)t(x) + A(1 - t(x)), \tag{4}$$

$$t(x) = e^{-\beta d(x)}, \tag{5}$$

where $J(x)$ is the scene radiance ("clean image"), A denotes the global atmospheric light, β is the scattering coefficient of the atmosphere, and $d(x)$ is the distance between the object and the camera.

The formula for deformation is:

$$J(x) = \frac{1}{t(x)}L' - A\frac{1}{t(x)} + A. \tag{6}$$

The two parameters $t(x)$ and A are unified into one formula, that is, $K(x)$ in the following formula, and the reconstruction error in the pixel domain is directly minimized. To this end, the formula is expressed as the following conversion formula:

$$J(x) = K(x)L' - K(x) + m, \tag{7}$$

where

$$K(x) = \frac{\frac{1}{t(x)}(L' - A) + (A - m)}{L' - 1}. \tag{8}$$

In this way, $t(x)$ and A are integrated into a new variable $K(x)$. m is a constant deviation with a default value of 1. Because $K(x)$ depends on L', the objective is to construct

an input adaptive depth model whose parameters will vary with the input low-light-level image to minimize the reconstruction error between the output $J(x)$ and the actual clear image on the ground. The LLAON consists of two modules. As shown in the Fig. 2(a), K estimation module is used to estimate $K(x)$ from input I, and then $K(x)$ is used as an adaptive parameter to input a clear image generation module to estimate $J(x)$ to obtain a clear image with improved contrast.

The K estimation module is a key component of the LLAON; it is responsible for estimating the depth and the relative light level. In the convolution network, the shallow features are more detailed, and there is more deep semantic information. The fusion of the shallow features and the deep features can increase the amount of useful information to a certain extent so that the K-map can be estimated more accurately. As shown in the Fig. 2(b), five convolution layers are used, and multi-scale features are formed by fusing convolution cores of different sizes. In [23], parallel convolution with different convolution core sizes is used in the second layer. Ref. [24] links the features of coarse-scale networks with the middle layers of fine-scale networks. Inspired by [23, 24], the characteristics of the "concat1" layer and connection layers "conv1" and "conv2" of the LLAON are linked. Similarly, the "concat2" connection comes from the characteristics of "conv2" and "conv3"; the "concat3" connection comes from the characteristics of "conv1", "conv2", "conv3" and "conv4". This multi-scale design captures the characteristics of different scales, and the

intermediate connection also compensates for the information loss during convolution. Each convolution layer of the LLAON uses only three convolution cores. Therefore, the LLAON is lighter than existing deep-seated methods. After the K estimation module, the clear image generation module is composed of an element-by-element multiplication layer and several element addition layers so that the restored image can be generated by formula (6). The calculation flow of this algorithm is shown in Table 1.

3.4 Image quality evaluation

3.4.1 Peak signal–noise ratio (PSNR)

The PSNR is an objective criterion for image evaluation. It measures the damage degree of the original image and the noise and is similar to a human perception of the image. Additionally, the higher the PSNR is, the better the denoising image. Basically, it is a correction of the mean-square error between the clear-light reference image and the reconstructed image. Given a clear-light $m \times n$ image I and its reconstructed version K, the MSE is calculated by the formula:

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=1}^{n-1} [I(i,j) - K(i,j)]^2, \tag{9}$$

and the PSNR, in dB, is expressed as:
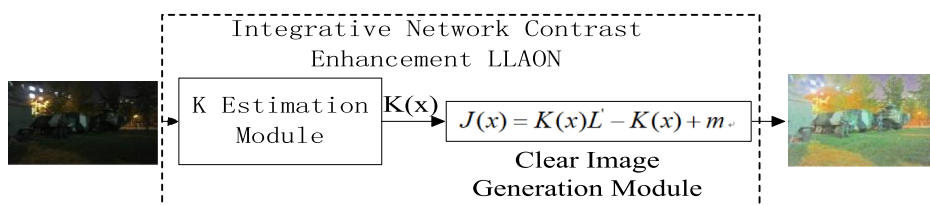
(a)
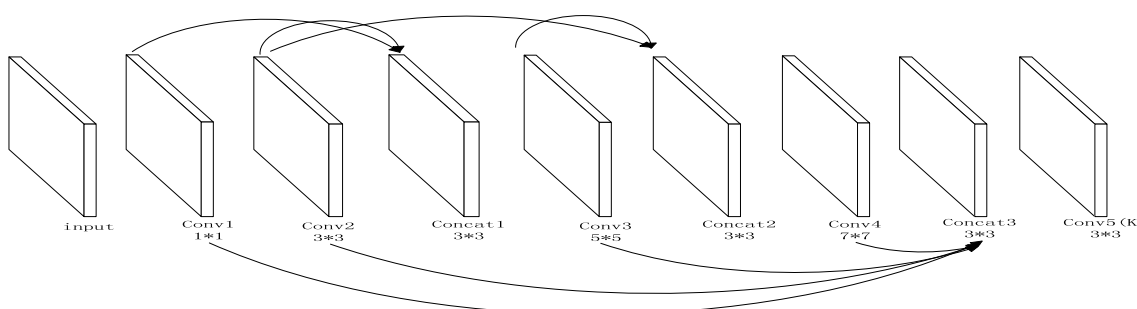
(b)

Fig. 2 **a** Integrated network model. **b** K estimation module

Table 1 Algorithmic flow chart

The algorithm in this paper: low-light-level image enhancement based on an integral network
Input: a low-light-level image I
1: Super-pixel segmentation is applied to I, and λ_i is calculated by formula (1)
2: Image I is reversed, and a new image L is obtained after inversion
3: BM3D filters are used to obtain $b^{\text{fine}}(L)$ and $b^{\text{coarse}}(L)$, and $b(L)$ is obtained by combining these formulas
4: The first-order difference method is applied to L' to generate the noise detail layer $d_1(L)$
5: Structural filter smoothing $d_1(L)$ is used to generate the noise-free detail layer $d(L)$
6: First, the noise-free detail layer $d(L)$ is adaptively combined with the basic layer $b(L)$, and then L is obtained according to formula (2)
7: Through the K estimation module in the integrated network, $K(x)$ is obtained
8: $J(x)$ is estimated by the clear image generation module
9: Image $J(x)$ is reversed to obtain the final output enhancement image E
Output: An enhanced image E

$$\text{PSNR} = 10 \cdot \log_{10}\left(\frac{\text{MAX}_I^2}{\text{MSE}}\right), \tag{10}$$

where MAX_I is the maximum possible pixel value of image I. Because the pixels used in this paper are 8 bits, in this case, $\text{MAX}_I = 2^8 - 1 = 255$.

3.4.2 Structural similarity (SSIM)

The structural similarity index (SSIM) is used to measure the structural similarity between two images and is sensitive to local structural changes. Under specific illumination conditions, image enhancement will not change the structure, and the texture of the images remains relatively complete. For example, in natural images taken during the day, all the pixel values can be changed by adding or subtracting a small number, but the structure and texture will not greatly change, while the PSNR behaves quite differently. For low-light-level image enhancement, texture preservation is more important than structure preservation. Therefore, the SSIM function is more suitable for low-light-level image enhancement than other kinds of enhancement. The SSIM is equivalent to calculating the illumination (the mean value of image sub-regions), contrast (variance of image sub-regions) and normalized pixel vectors after normalizing the data, and then multiplying the values.

The SSIM value of pixel i is calculated by definition as:

$$\text{SSIM}(i) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \cdot \frac{2\delta_{xy} + C_2}{\delta_x^2 + \delta_y^2 + C_2}. \tag{11}$$

The SSIM value is in the range (0,1), and the value 1 means that the two images are identical. Therefore, $1 - \text{SSIM}(i)$ is used to calculate the loss of pixels. The SSIM function is defined as follows:

$$l_{\text{SSIM}} = \frac{1}{N}\sum_{i \in I} 1 - \text{SSIM}(i). \tag{12}$$

We divide image I into N sub-regions, where i is a sub-region of region I.

4 Experimental results

4.1 Introduction of the experimental data set

Because deep learning requires a large number of training samples, it is difficult to achieve complete consistency between a low-illumination image taken under a low light level and a normal-illumination image of the same scene, which seriously restricts the research and development of deep learning technology in low-light-level image enhancement. Inspired by [25], this paper proposes an easy-to-operate but time-consuming training sample generation method. According to the retinex model, a low-illumination image is the product of the illumination component M and the reflection component N (the normal-illumination image). Experiments are carried out on the Berkeley Segmentation Data Set (BSD) in the field of computer vision. Five hundred images with good illumination conditions are selected as reflective components. A total of 256,000 image blocks with dimensions of 40 pixels by 40 pixels are randomly selected and fused with randomly selected images with Gaussian white noise to obtain the illumination component L. The fusion ratio $\beta, \beta : [0, 1]$ obeys a beta distribution. Each pixel corresponding to the two images is added. The value is added directly—that is, $L = \beta * \text{image} + (1 - \beta) * \text{image\_random}$ —and the synthesized low-illumination image block can be expressed as $S = M*N$. The data can be expanded to 26,000 low-light-level images with a size of 480*640 by the above transformation and a scale transformation. To make the data more diversified, the relevant video images on the network are captured, the data sets are constructed, and 4000 images are captured. After the expansion of the original data and the image on the network, the data set has a total of 30,000 pictures. The training set of this method is constructed as follows: 70% are the data extracted randomly from the data
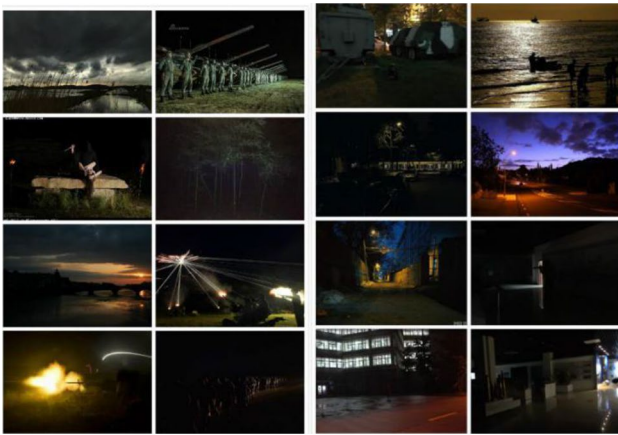
Fig. 3 Partial data set

set, and 30% are the data extracted from the data set. Each picture is labeled manually according to the format of the BSD data set. Some of the data set pictures are shown in Fig. 3.

4.2 Experimental settings

In the Ubuntu 16.04 operating system, a low-light-level image enhancement algorithm based on an integrated network is implemented using Pytorch, a deep learning framework. The experimental platform uses an Intel (R) Core (TM) i5-8600 central processing unit (CPU), and its main frequency is 3.10 GHz; it has 16-G memory, and its GPU (graphics processing unit) is an NVIDIA GTX 1080Ti. We train and test the network in the above environment. To verify the effectiveness and optimization of the algorithm, seven current algorithms with good performance are selected for comparison under the same environment. Regarding the training set for this method, 70% of them are data extracted randomly from the data set, and 30% are the data extracted from the data set. Additionally, real pictures are used for testing.

The algorithm network has five convolution layers, each of which has a size of 2.3 sections. The number of convolution kernels in the feature extraction part of the network framework is 64, and the size is 3*3*64 pixels; the number of all convolution kernels in the non-linear mapping part is 64, and the size is 5*5*64 pixels; the convolution kernels in all layers of the network are adopted. In the initialization method of [26], in the training process, the weights are initialized by using Gaussian random variables, the bias term is initialized to 0, the momentum and attenuation parameters are set to 0.9 and 0.0001, respectively, and the learning rate is 0.01. In the later stage, in the process of fine-tuning and retraining, over fitting is prevented, and the learning rate is reduced by an order of magnitude to perform retraining. The

size, C_1 and C_2 parameters of the SSIM kernel are set to 8, 0.001 and 0. 0001; 10 000 iterations are carried out, and a simple mean-square error (MSE) loss function is used. The experiment shows that not only the PSNR but also the SSIM and visual quality are improved.

4.3 Experimental analysis

In this section, through objective quantitative analysis and subjective visual analysis, the effectiveness and optimization of the algorithm are analyzed. The adopted image size is consistent, and the full reference mode is selected to evaluate the image quality. Next, the proposed model is objectively compared with other traditional contrast enhancement methods and depth-based learning methods. Then, the proposed algorithm is analyzed on the basis of the objective vision and compared with other algorithms in terms of visual performance.

4.3.1 Objective quantitative analysis

To validate the effectiveness and optimization of the algorithm, six enhancement methods are selected and compared under the same test environment: classical histogram equalization (HE), dynamic histogram equalization (DHE), Dong's image enhancement method based on the defogging model, LIME, a CNN, an LLCNN based on deep learning and the method proposed in this paper [27–29]. The results are shown in Table 2.

In the objective evaluation, Table 2 is the average of some evaluation indexes obtained on the same data set using different enhancement algorithms. The bold part are the performance results of the proposed algorithm. We compared the PSNR, SSIM, MSE and LOE (lightness order error); a higher PSNR reflects a smaller distortion degree of the image, a higher SSIM represents better integrity of the image structure information, and the MSE reflects the structural characteristics of the enhanced image and the real image. The difference between the enhanced image and the real image is reduced; the brighter the enhanced image is, the closer it is to the original normal-illumination image.

Table 2 Evaluation indicators for different enhancement algorithms

METHOD	PSNR/DB	SSIM	MSE	LOE
HE	16.07	0.7885	1928.3	62.68
DHE	16.35	0.6469	1899.8	58.49
DONG	17.29	0.7848	1684.1	95.65
LIME	19.17	**0.9251**	1303.7	93.60
CNN	23.08	0.790	998.6	36.68
LLCNN	30.20	0.853	586.4	29.27
Ours	**31.64**	**0.912**	**496.7**	**25.36**

The LOE mainly evaluates the naturalness preservation ability of the enhanced image. The smaller the value, the more the brightness sequence is protected and the more natural the image is. Their formulas are as follows:

MSE: MSE is the energy mean of the difference between the real image and noisy image.

$$\text{MSE} = E[(y_b - y_a)^2], \tag{13}$$

where y_b represents the real image and y_a is the noisy image.

PSNR: PSNR is the ratio of the peak signal energy to the MSE.

$$\text{PSNR} = 10 \log_{10} \frac{\text{MaxValue}^2}{\text{MSE}} = 10 \log_{10} \frac{2^{\text{bits}} - 1}{\text{MSE}}. \tag{14}$$

In the second equation, "bits" is the number of bits that each pixel holds; the pixel value of the image is stored quantitatively. Thus, MaxValue is $2^{\text{bits}} - 1$.

SSIM: The brightness is represented by the mean μ, the contrast by the variance normalized by the mean δ, and the structure by the correlation coefficient C.

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\delta_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\delta_x^2 + \delta_y^2 + C_2)}. \tag{15}$$

LOE: The relative brightness order is used to measure the naturalness degree of the image. The order of relative brightness can be used to indicate the direction of illumination and the degree of the variation of illumination.

The LOE measures the order of difference in brightness between enhanced pictures I_e and I:

$$L(x, y) = \max_{c \in \{r,g,b\}} I^c(x, y). \tag{16}$$

$L(x, y)$ is the maximum value in the RGB channel. For each pixel, the relative brightness order difference between the original image and the enhanced image is defined as:

$$\text{RD}(x, y) = \sum_{i=1}^{m} \sum_{j=1}^{n} (U(L(x, y), L(i, j)) \oplus U(L_e(x, y), L_e(i, j))), \tag{17}$$

$$U(x, y) = \begin{cases} 1, & \text{for } x \geq y \\ 0, & \text{else} \end{cases}, \tag{18}$$

where $\oplus$ is the xor operation.

$$\text{LOE} = \frac{1}{m * n} \sum_{i=1}^{m} \sum_{j=1}^{n} \text{RD}(i, j). \tag{19}$$

From Table 2, we can see that, in addition to the SSIM being slightly lower than that of LIME, the PSNR, MSE and LOE are better than those of the other algorithms. This shows that the proposed algorithm has low distortion and good texture details, is closer to the original image, and

produces an enhanced image that is more realistic and natural. The effectiveness and optimization of the algorithm are verified.

Different methods are compared in terms of the time cost. To ensure a comparable performance in terms of the image processing time, in the same hardware environment, we pre-process the data of the following algorithms, fine-tune the network, and ensure a time-consuming comparative analysis under the same image size (244*244). The experimental results are shown in Fig. 4. Although LIME causes very little distortion, it is quite time-consuming. The proposed method achieves less distortion than the other methods at an acceptable time cost.

From Fig. 4, we can see that the classical low-light-level image contrast enhancement algorithm has a simple structure. For example, the HE, DHE, DONG, and lIME algorithms mainly process the input pixels directly, and process them quickly through functions and formulas, so their time costs are low. The low-light-level image enhancement method based on depth learning needs to perform parameter adjustment because of its complex structure, and deep learning processes a large number of image features into the pixel matrix, including many convolution operations, leading to a large time consumption. The CNN algorithm uses many convolution layers for processing, and the time cost is large, while LLCNN uses residual blocks later, which further increases the time cost. Because the integrated network parameters mentioned in the method of this paper are fewer, the number of network layers is smaller and the nested formulas function better as a network compared with other methods. Based on the convolution neural network enhancement algorithm, this algorithm is lighter, less time consuming and performs faster image processing.

4.3.2 Subjective visual analysis

In subjective evaluation, all methods are tested on 32 real images with different darkness and noise levels. Due to
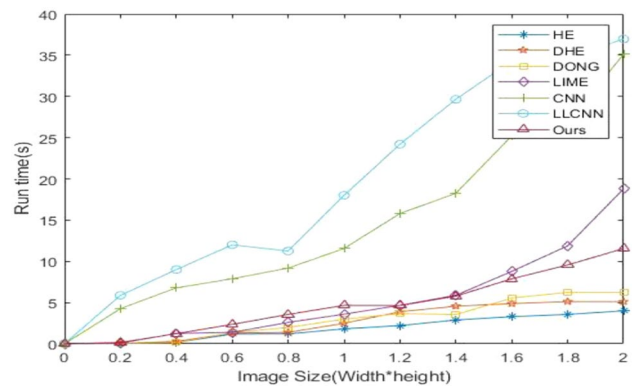

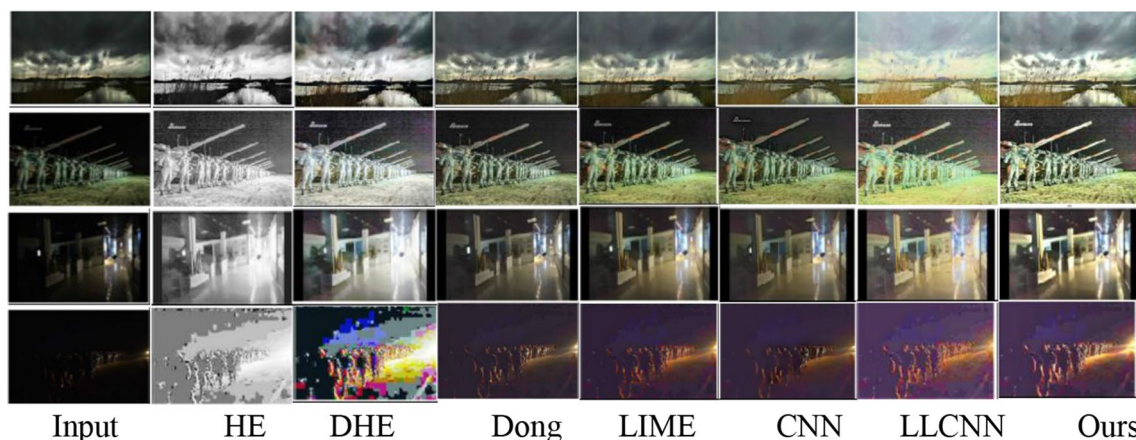**Fig. 4** Time consumption of different algorithms

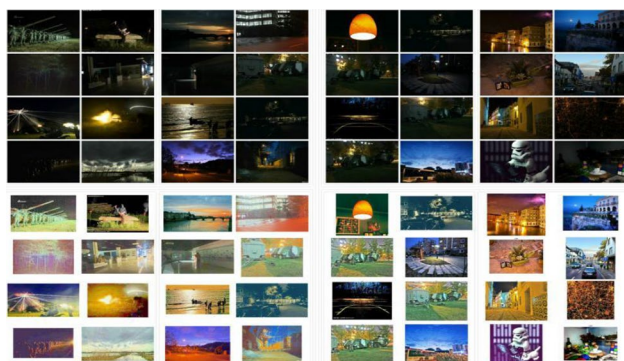**Fig. 5** Comparison of contrast enhancement algorithms for different low-light-level images

Fig. 6 32 real test images and enhanced results

Fig. 7 Average LOE comparison of 32 images with different algorithms

space limitations, this section compares the experimental results of four images with different darkness and noise levels using different algorithms. The sizes of the input image are 512*512 and 1024*1024. For different algorithms, corresponding pre-processing is added, and the normalized image size is 512*512 pixels so that the final output image is a fixed size and the experimental comparison effect is more obvious; the images obtained by other enhancement algorithms are then compared. The results show that the proposed algorithm can adaptively control the contrast enhancement of different regions to prevent over-enhancement and suppress most of the noise while retaining the texture details. Figure 5 shows the comparison results. Figure 6 shows 32 images and the enhancement results obtained by this algorithm, and Fig. 7 shows the plane of the 32 images. All LOEs were compared.

Figure 5 shows that although the HE algorithm can enlarge the dynamic range of the image and improve the overall brightness and contrast of the image, over-saturation will lead to serious color bias. From the image, it can be clearly observed that the sky/indoor color is white,
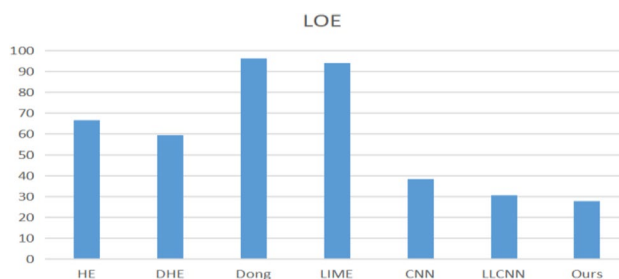
while that in the original image is black, the grass color is no longer green, and the overall image is yellow. Similar color distortion also occurs in DHE. The Dong algorithm maintains the image color well after enhancement, but it has a serious edge effect. In the figure, observing soldiers, plants and construction equipment shows that the edges of these elements have obvious black lines, and the image looks very unnatural. The LIME algorithm improves the brightness of the dark areas, improves the overall brightness of the image, and makes the color brighter. However, the LIME algorithm is prone to over-enhancement and causes some color distortion. The CNN algorithm is only effective for dark areas, but it cannot improve the brightness of dark areas, and it is difficult to distinguish the details clearly. The brightness of human and indoor equipment images is still dark after enhancement. The LLCNN algorithm greatly enhances the brightness of low-light-level images, but it is easy for it to over-enhance. Soldiers in the figure are revealed, and the LLCNN algorithm will enhance an area whether the area is dark or not. Therefore, the image is over-enhanced, which is not in line with human visual perception. The proposed algorithm can maintain the color and improve over-enhanced images,

enhance the brightness of dark areas, and have a better visual perception.

For all 32 images in Fig. 6, the average LOE values of seven different methods of low-light-level image enhancement are calculated in Fig. 7. The results show that the algorithm presented in this paper has better naturalness preservation ability on the basis of retaining details and textures.

5 Conclusion

Aiming at addressing the phenomena of color distortion and over-enhancement that arise when the current classic low-light-level image enhancement algorithm improves the brightness and contrast of an image, an effective integrated neural network algorithm is proposed to enhance the contrast of low-light-level images. Using adaptive de-noising based on super-pixels and adaptive contrast enhancement based on an integrated neural network, the heavy noise and texture blur in traditional methods are eliminated, and the contrast of low-light-level images is effectively improved. The experimental results show that the proposed algorithm not only improves brightness and contrast but also avoids serious color distortion. The enhancement effect is better than that of the current mainstream low-light-level image enhancement algorithm, which has theoretical significance. In the future, we will continue to optimize the network model to further improve the speed and performance of image processing.

References

1. Talaulikar, A.S., Sanathanan, S., Modi, C.N.: An enhanced approach for detecting helmet on motorcyclists using image processing and machine learning techniques. Advanced Computing and Communication Technologies (2019)
2. Yan, C.: Several contrast enhancement algorithms for low-light images [J]. J. Jingdezhen Univ. **6**, 10–12 (2013)
3. Bo, P., Yiming, W., Pengbo, et al.: Research and implementation of low illumination image enhancement algorithm. Comput. Appl. (6) (2007)
4. Lili, D., Chang, D., Wenhai, X.: Two improved methods of image enhancement based on histogram equalization [J]. J. Electron. **46**(10), 65–73 (2018)
5. Xueming, L.: Image enhancement algorithm based on Retinex theory [J]. Comput. Appl. Res. **22**(2), 235–237 (2005)
6. Li, L., Wang, R., Wang, W., et al. A low-light image enhancement method for both de-noising and contrast enlarging[C]//2015 IEEE International Conference on Image Processing (ICIP). IEEE, September 27–30, 2015, Quebec City, Canada
7. Yandong, L., Zongbo, H., Leihang. Review of convolutional neural networks [J].Comput. Appl. **36**(9), 2508–2515 (2016)
8. Hou, J.-C., Wang, S.-S., Lai, Y.-H., et al.: Audio-visual speech enhancement based on multimodal deep convolutional neural network (2017). arXiv preprint arXiv:1804.03641
9. Chao, C., Chao, C. Application of improved single-scale Retinex algorithm in image enhancement [J]. Comput. Appl. Softw. **30**(4), 55–57 (2013)
10. Watanabe, T., Kuwahara, Y,. Kojima, A., Kurosawa, T.: Improvement of color quality with modified linear multiscale retinex. Proc. SPIE 5008, Color imaging VIII: processing, hardcopy, and applications, 13 January 2003 (2003). https://doi.org/10.1117/12.472030
11. Chun, H., Guoqin, S.: Algorithmic implementation based on MSRCR theory [J]. Intell. Comput. Appl. **7**(3), 114–116 (2017)
12. Fu, X., Zeng, D., Yue, H. et al. A weighted variational model for simultaneous reflectance and illumination estimation[C]//computer vision & pattern recognition. Las Vegas, Nevada, USA, Jun 26, 2016–Jul 1, 2016
13. Guo, X., Li, Y., Ling, H.: LIME: low-light image enhancement via illumination map estimation[J]. IEEE Trans. Image Process. **26**(2), 982–993 (2017)
14. Dong, C., Loy, C.C., He, K., et al.: Image super-resolution using deep convolutional networks[J]. IEEE Trans. Pattern Anal. Mach. Intell. **38**(2), 295–307 (2014)
15. Guoliang, Y., Nan, X., Fang, L., et al. Research on the deep learning classification method of non-linear activation function [J]. J. Jiangxi Univ. Technol. **39**(193)(3), 79–86 (2018)
16. Yandong, L., Zongbo, H., Leihang. A review of Convolutional neural network [J]. Comput. Appl. **36**(9), 2508–2515 (2016)
17. Djurović, I.: BM3D filter in salt-and-pepper noise removal[J]. Eurasip J. Image Video Process. **2016**(1), 13 (2016)
18. Huanqing, Q.: Design of lattice structure filter [J]. Dig. Technol. Appl. **5**, 181–185 (2014)
19. Abdullah, H.R., Aleawy Algazi, N.A., Bayes estimators for the shape parameter of pareto type i distribution under generalized square error loss function[J]. Soc. Sci. Electron. Publ. 20–32 (2014)
20. Xiaoqiao, H., Junsheng, S., Jian, Y., et al. Evaluation of color image quality based on mean square error and peak signal-to-noise ratio [J]. J. Photon. **36**(Sup1)
21. Tong, Y.B., Zhang, Q.S., Yun-Ping, Q.I.: Image quality assessing by combining PSNR with SSIM[J]. J. Image Graph. **11**(12), 1758–1763 (2006)
22. Bergh, M.V.D, Boix, X., Roig, G., et al. SEEDS: superpixels extracted via energy-driven sampling[C]//European conference on computer vision. Springer, Berlin (2012)
23. Cai, B., Xu, X., Jia, K., et al. DehazeNet: an end-to-end system for single image haze removal.[J]. IEEE Trans. Image Process. **25**(11), 5187–5198 (2016)
24. Ren, W., Si, L., Hua, Z., et al. Single image dehazing via multi-scale convolutional neural networks[M]//computer vision—ECCV 2016. 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016
25. Zhang, Z, He, T, Zhang, H, et al. Bag of freebies for training object detection neural networks (2019). arXiv preprint arXiv:1902.04103
26. He, K., Zhang, X., Ren, S., et al. Delving deep into rectifiers: surpassing human-level performance on imagenet classification (2015). arXiv preprint arXiv:1502.01852
27. Weihe, Z., Junwei, Yang, Yuxiang, X.: Bidirectional equalization of dynamic histogram for image enhancement [J]. Stereol. Image Anal. China **2**, 129–139 (2014)
28. Hongqiang, M., Shiping, M., Yueli, X., et al. Low illumination image enhancement based on deep convolution neural network [J]. J. Opt. **39**(2) (2019)
29. Li, T., Zhu, C., Xiang, G., et al. LLCNN: a convolutional neural network for low-light image enhancement[c]//visual

communications & image processing. Taichung, Taiwan, December 09 (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Peng Wang received a B.S. degree from Xi'an Technological University in 2010. He is currently a professor in the department of Electronic Information Engineering at Xi'an Technological University. His research interests include homeland security, sensors and sensor networks, and artificial intelligence.

Jiao Wu is working towards a B.S. degree in the Department of Communication Engineering at Xi'an Technological University. Her main research directions are machine vision and deep learning.

Haiyan Wang received a Ph.D. degree from Northwestern Polytechnical University in 2004. He is currently a professor in the department of Electronic Information and Artificial Intelligence at Shaanxi University of Science and Technology. His research interests include underwater acoustic communication and networking, target recognition and location tracking.

Xiaoyan Li received a Ph.D. degree in Communications and Information Systems from Xidian University in 2013. She is currently an associate professor in the department of Electronic Information Engineering at Xi'an Technological University. She is currently working on spectrum sensing and MAC protocol design in cognitive radio networks.

Yongxia Yang received a B.S. degree from Xi'an Technological University in 1996. She is currently a professor in the department of Electronic Information Engineering at Xi'an Technological University. She has over 40 years of experience in electromagnetic interference. Her expertise encompasses theoretical physics, sensor networks and artificial intelligence.