Ukrainian Journal of Physical Optics
ISSUE 1
2025

LOW-LIGHT IMAGE ENHANCEMENT BASED ON DEPTHWISE SEPARABLE CONVOLUTION

Y. QIU [1], H. WANG [2], T. DEMKIV [3*], O. KOCHAN [2,4], L. YAN [2]

[1] Department of Computer, Hubei University of Technology Engineering and Technology college, Wuhan 430068, China
[2] School of Computer Science, Hubei University of Technology, Wuhan 430068, China
[3] General Physics Department, Ivan Franko National University of Lviv, 8 Kyryla and Mefodiya Str., 79005 Lviv, Ukraine
[4] Department of Measuring-Information Technologies, Lviv Polytechnic National University, 79013 Lviv, Ukraine
*Corresponding authors: tmdemkiv@gmail.com

Received: 23.09.2024

Abstract. Since the appearance of deep learning algorithms, convolution neural networks (CNN) based algorithms have significantly progressed in weak light image enhancement. However, they still face a major problem: the CNN-based low illumination enhancement algorithm has excessive computational complexity and needs sufficient memory. Although the algorithm's accuracy is improved, the computational efficiency is reduced. This paper introduces the lightweight network for low illumination, and image enhancement is proposed. We first introduce the background of the technology used. Based on the principle of MobileNetV2, we use the generative adversarial networks with improved attention mechanisms as our base algorithm. Then, three comparative algorithms are built for experiments. Experiment results confirm that the proposed network needs fewer algorithm parameters while guaranteeing the low-light image enhancement effect.

Keywords: machine learning, image enhancement, adversarial networks, depthwise separable convolution
UDC: 004.89
DOI: 10.3116/16091833/Ukr.J.Phys.Opt.2025.01040

1. Introduction

Researchers have paid more attention to image enhancement as a hot research direction in image processing. Many image processing techniques were proposed to improve the quality of low-light images effectively. Physics-based and deep learning-based algorithms can be categorized depending on the algorithm used.

Of course, the physics-based algorithm came up first. Data processing is key in improving instruments [1, 2]. Statistical methods were first adapted to enhance the quality of images. The histogram equalization (HE) [6] is a widely used method for enhancing an image's contrast. This method is simple to implement and has high computational efficiency, so it has been widely used. However, due to the serious noise and low brightness of low-illumination images, this method usually cannot take into account brightness, texture details, and color, so it still has some problems, such as local overexposure. Since low-lighting images are usually distorted by Poisson noise with signal-to-noise ratio of 10-20 dBs [7-9], demosaicing techniques applied in [7-9] in which the experiments showed that more white pixels are the better they work. The Retinex algorithm [10] effectively adjusts brightness and improves an image's colors. It stimulates the imaging principle of the human retinal cortex, decomposes the image into a light and reflection component, and builds the corresponding

algorithm. To effectively estimate these two components, eliminate the illumination component in the image, and propose the reflection component to enhance the function of low-light images, researchers proposed the single-scale Retinex theory [11], multi-scale Retinex theory (MSR) [12], and other methods. Fu et al. [13] obtained the illumination component using the adaptive Gaussian filter, eliminated the artifacts, and restored it using the brightness adaptive method. Experiments proved that this method effectively improved the visual appearance of low-light images. Image enhancement methods are also based on filtering [14], especially the wavelet transform, which has gained extensive acceptance in image processing. In 2009, He Kaiming et al. [15] proposed the definition of dark channel priors in mathematics, yielding a simple and efficient algorithm that has been widely used. Li et al. [16] reduced the estimation error and running time by improving the estimation and transmittance of atmospheric illumination and introduced a four-dimensional binary tree to avoid a halo in the enhanced image.

Since deep learning was introduced into the field of image enhancement, the method based on deep learning has gradually replaced the physical algorithm-based low-light image enhancement method. The following are some representative literature in this field. Ref. [17] is an early paper that uses deep learning to complete low-light enhancement tasks. It supports the stacking sparse denoising autoencoder based on synthetic data training. It can enhance and de-noise low-illumination noisy images. In [18], CNN was introduced, and the conventional MSR method can be considered a feedforward convolutional neural network with different Gaussian convolution kernels. In fact, the paper [19] is centered on single-image contrast enhancement, addressing the issue of low contrast in underexposed and overexposed situations. Ref. [20] utilized a two-stage step decomposition followed by enhancement, which the CNN entirely implemented. The key point of the paper [21] is the extraction and fusion of features at different levels in the network. Ref. [22] focuses on image enhancement under extremely low-light and short-time exposure conditions. Similarly, many applications in reality also use the abovementioned theory. In paper [23], intelligent systems execute various tasks such as object detection using enhanced images, classification, segmentation, recognition, scene understanding, and 3D reconstruction in the Smart City concept. Paper [24] introduces the method to solve the ambiguity of pictures taken by a camera. Image enhancement technology has wide applications in other areas, such as mobile phone shooting, criminal investigations, digital cameras, etc. [25, 26].

As deep learning algorithms evolve, this paper considers depthwise separable convolution and its adaptation to low-light image enhancement tasks in the spatial domain. Inspired by the lightweight network [27] series such as MobileNet and Shufflenet, we improved the lightweight and low-light image enhancement network. First of all, according to the conventional depth of the separable convolution module, we remove the first rectified linear unit (ReLU) activation function; the aim is to avoid the ReLU function loss caused characteristics half the probability of 0, at the same time removing all batch regularization, add an instance of regularization [28]. Instance regularization is the operation of removing instance-specific contrast information from content images, which prevents mean and covariance shifts and simplifies the learning process, yielding an improved depth separable convolution module (IN–DepthwiseConv). Second, to facilitate the componentization of the depth separable convolution module, we proposed an improved inverse residual depth

separable convolution module (IN-Bottleneck) based on the residual depth of the separable convolution module. Next, we build three algorithms based on IN-DepthwiseConv and In-Block named DwDG, DwG, and DeeperDwG algorithms, respectively. The DwDG algorithm uses IN-DepthwiseConv and the IN-Bottleneck to replace the corresponding convolution layer in the generator and discriminator. The DwG algorithm only replaces the generator layer with the IN-DepthwiseConv module. The DeeperDwG algorithm replaced the generator with the deeper IN-Bottleneck module. Finally, the experimental comparison of the proposed algorithm with several known algorithms shows that the proposed one can decrease the number of algorithm parameters and speed up the process while keeping the enhancement loss effect small. The experimental comparison was carried out for the synthetic and the real-world data sets.

The following points are major contributions of this article:

1. We use a lightweight network to improve the low-light image enhancement network, improving the operating efficiency of the network.
2. We apply the depthwise separable convolution module to improve the ordinary convolution layer in the network, which improves the enhancement effect on low-light images.
3. Experiments on real datasets confirm our proposed algorithm's performance and that the proposed one fully meets the application requirements of trusted multimedia computing in urban scenarios.

The structure of this article is as follows:

In Section 2, the related works are considered. In Section 3, the principle of the proposed algorithm will be explained. In Section 4, we conducted comparative experiments to demonstrate the superiority of the proposed algorithm. Finally, in Section 5, we conclude the article and announce future work.

2. Related works

This section introduces the techniques used to build the improved depthwise separable convolution generative adversarial networks algorithm, including deep separable convolution and the lightweight network.

2.1. Depthwise separable convolution

In the standard convolution process, all channels in the image region are considered simultaneously, and convolution is performed in the same way. Depthwise separable convolution presents a novel point: different input channels use different convolution kernels. It decomposes an ordinary convolution operation into two processes: deep convolution and point-by-point convolution.

Fig. 1a represents the standard convolution. Assuming the size of the input feature graph is $D_K \times D_K \times M$, the size of the convolution kernel is $D_K \times D_K \times M$, and the size of the output feature graph is $D_K \times D_K \times N$, then the number of parameters of the standard convolution layer is $(D_K \times D_K \times M) \times N$, where D_K is the spatial dimension of the kernel assumed to be square, M is the number of input channels, and N is the number of outputs.

Fig. 1b represents the depthwise convolution, which acts as a filter. The size of it is $(D_K, D_K, 1)$. It has M convolution layers acting on each channel of the input, and the number of parameters is $(D_K \times D_K \times 1) \times M$. Fig. 1c represents Pointwise convolution with the size of $(1,1,M)$, a total of N convolution points, acting on the output feature mapping of deep convolution, and the number of parameters is $(1 \times 1 \times M) \times N$.

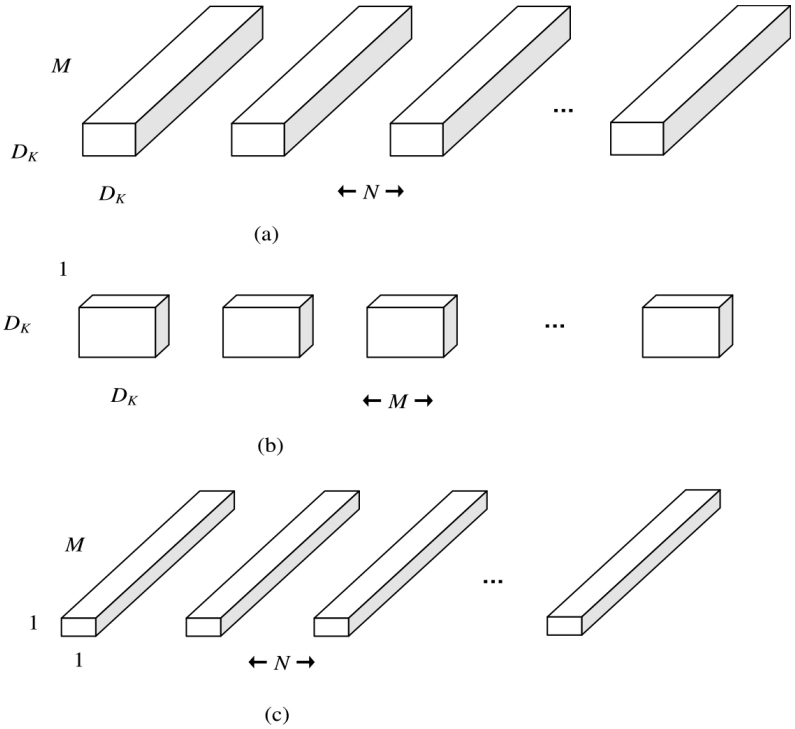

Fig. 1. Schematic diagram of the standard convolution and depthwise separable convolution: (a) standard convolution; (b) depthwise convolution; (c) pointwise convolution.

The combination of depthwise convolution and 1×1 convolution is depthwise separable convolution. Combined with the above formulae, it can be shown that the number of deep separable convolution parameters is:

$$\frac{D_K \times D_K \times M + M \times N}{D_K \times D_K \times M \times N} = \frac{1}{N} + \frac{1}{D_K^2}. \tag{1}$$

To sum up, if only extracting an attribute, the method of depthwise separable convolution is more inefficient than the standard convolution. Depthwise separable convolution can save more parameters by extracting increasing attributes.

2.2. Lightweight network

As computer vision research develops based on deep CNN, many applications of computer vision scenes in reality are driven, such as face recognition, ID card character recognition, etc. With the development of Internet technology, mobile devices have gradually replaced personal computers and become an important part of people's daily lives. In such an environment, the application and service of artificial intelligence began to turn their attention to cell phones. However, the lightweight deep CNN for mobile terminals was invented due to the limitation of storage space and computing capacity of mobile terminals. This is where the lightweight web comes from.

Next, we consider two typical lightweight networks to explain how they work.

The core of MobileNet is that it uses depthwise separable convolution. The core of depthwise separable convolution is the decomposition of standard convolution operations, depthwise convolution, and pointwise convolution. This method reduces the computational complexity and the number of parameters of the convolution operation so the deep

convolutional neural network can adapt to the limited computational and storage resources.

MobileNet [29] is a lightweight network structure put forward by Google in 2017, aiming at mobile and embedded devices, making full use of limited computing resources and storage resources to achieve the best performance of the algorithm to meet the needs of various visual applications. Table 1 shows the network structure for MobileNet.

Table 1. The network structure of MobileNet.

Style/Tag	Convolution kernel dimension	Input size
Conv/s2	3×3×3×32	224×224×3
Conv dw/s1	3×3×32dw	112×112×32
Conv/s1	1×1×32×64	112×112×32
Conv dw/s1	3×3×32dw	112×112×64
Conv/s1	1×1×64×128	56×5×64
Conv dw/s1	3×3×128dw	56×56×128
Conv/s1	1×1×128×128	56×56×128
Conv dw/s2	3×3×128dw	56×56×128
Conv/s1	1×1×128×256	28×28×128
Conv dw/s1	3×3×256dw	28×28×128
Conv/s1	1×1×256×256	28×28×128
Conv dw/s2	3×3×256dw	28×28×128
Conv/s1	1×1×256×512	14×14×128
$5\times\begin{matrix}\text{Convdw / s1}\\\hline\text{Conv / s1}\end{matrix}$ Conv	$\frac{3\times3\times612\text{dw}}{1\times1\times512\times512}$	$\frac{14\times14\times512}{14\times14\times512}$
dw/s2	3×3×512dw	14×14×512
Conv/s1	1×1×512×1024	7×7×512
Conv dw/s2	3×3×1024dw	7×7×1024
Conv/s1	1×1×1024×1024	7×7×1024
AvgPool/s1	Pool×7	7×7×1024
Fc/s1	1024×1000	1×1×1024
Softmax/s2	Classifier	1×1×1024

Based on MobileNet, MobileNetV2 [30] improves the depthwise separable convolution module to a certain extent and proposes the reverse residual module of linear bottleneck, which makes MobileNetV2 better in both accuracy and efficiency compared with MobileNet. Table 2 shows the network structure of MobileNetV2.

The improvement points are as follows: 1) A linear activation function replaces the ReLU activation function in the output part of the depthwise separable convolution module to reduce information loss caused by ReLU calculation of a low-dimension tensor; 2). Adding the 1×1 point-by-point convolution operation before the deep separable convolution module to increase the number of channels of the input tensor, specifically, increase the dimension of the input tensor; 3). Using the jump connection to transfer the module's input to its output to combine with the output data (this step is not required for the module with a step size of 2). Fig. 2 shows the reverse residual module structure of a linear bottleneck.

Table 2. The network structure for MobileNetV2 algorithm.

Input	Operation type	Extension	Channel number	Operation number	Step Size
2243×3	Conv2d	-	32	1	2
1122×3	Bottleneck	1	16	1	1
1122×16	Bottleneck	6	24	2	2
562×24	Bottleneck	6	32	3	2
282×32	Bottleneck	6	64	4	2
142×64	Bottleneck	6	96	3	1
142×96	Bottleneck	6	160	3	2
72×160	Bottleneck	6	320	1	1
72×320	Conv2d 1×1	-	1280	1	1
72×1280	Avgpool 7×7	-	-	1	-
1×1×1280	Conv2d 1×1	-	K	-	

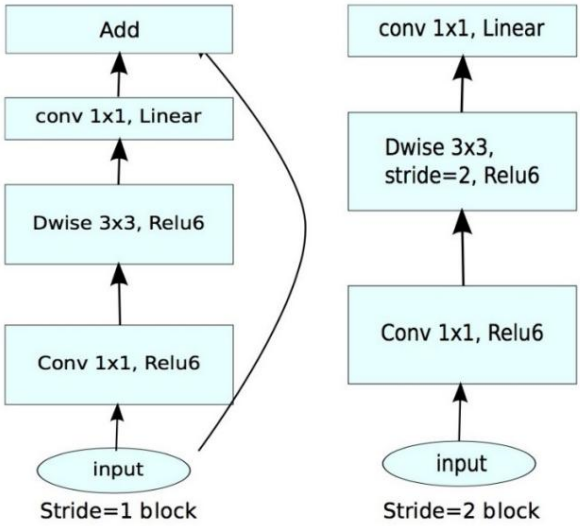

Fig. 2. The reverse residual module structure of a linear bottleneck [30].

The ShuffleNet [31] is a lightweight network algorithm of efficient computing proposed by Zhang, which has wide applications in mobile devices with limited computing power.

The ShuffleNet's network structure has two key innovations. One is pointwise grouping convolution, and the other is channel random shuffling. The operation of grouping convolution mainly divides and groups the image features equally along the channel dimension, and the corresponding convolution kernel carries out the calculation of feature extraction for each group of feature images. In this way, the parallel computation of multiple convolutions can be realized, and the efficiency of calculations can be improved. The operation of channel random shuffling is mainly to randomly mix the feature graphs calculated by each group and then distribute them to the network layer of the next group convolution. In the structure of ShuffleNet, the combination of block convolution and channel random shuffling facilitates the formation of the basic module of ShuffleNet. Fig. 3 shows the structure of this network [31].

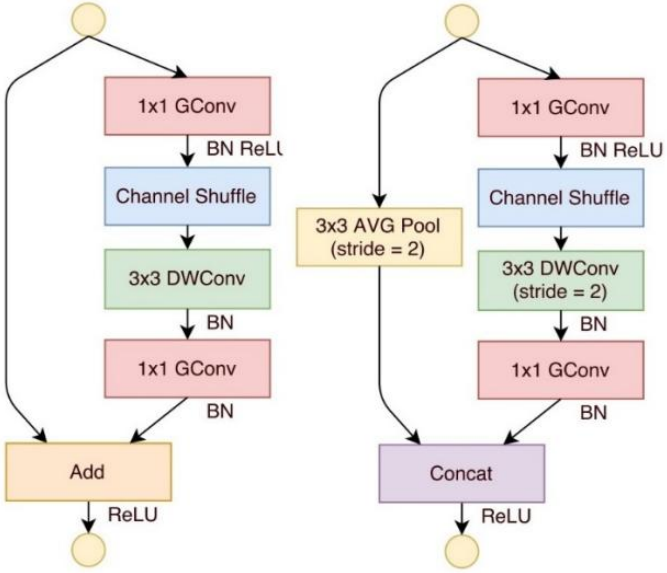

Fig. 3. The schematic diagram of the Shuffle Block structure.

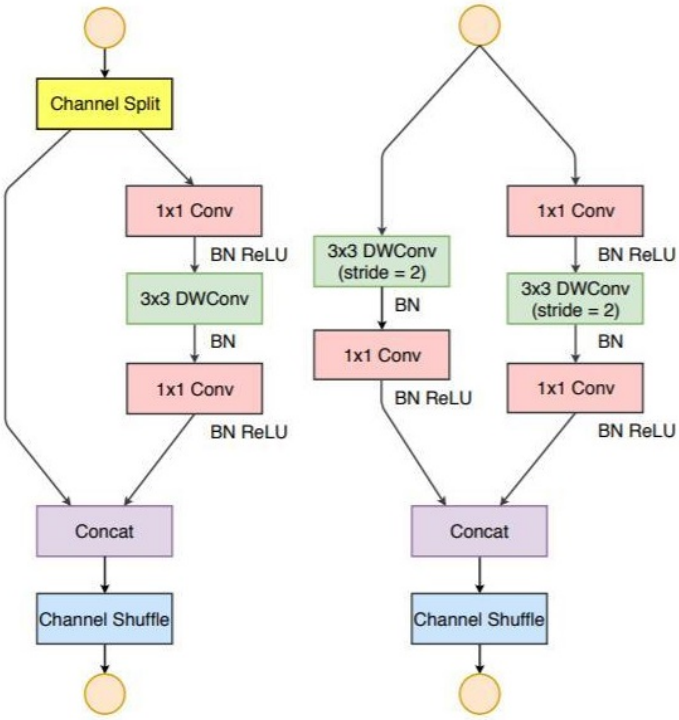

Fig. 4. Schematic diagram of the ShuffleNet v2 block.

ShuffleNet v2 was put forward in 2018 [32]. The algorithm is refined based on ShuffleNet. The proposed system for efficient guidance of the CNN should meet several requirements. First, the number of channels of the input and output characteristics influence the Memory Access Cost (MAC), hereinafter referred to as influence; only the same number of channels of the input and output characteristics ensures minimal MAC. Second, the grouping operations of the convolution layer greatly impact storage consumption. Excessive grouping operations will increase the storage consumption and slow the algorithm's running

speed. Therefore, grouping operations should be controlled. Third, the number of branches in the algorithm structure affects the algorithm's running speed. The fewer branches in the algorithm, the faster the algorithm speed. Fourth, the algorithm's speed is influenced by the element-by-element operation in the network, so this mode of operation will take a long time. Thus, the element-by-element operations should be reduced as much as possible. According to the proposed four requirements, ShuffleNet v2 is improved based on ShuffleNet, and the improvement is mainly for ShuffleBlock. Fig. 4 shows the block structure of ShuffleNet v2 [32].

2.3. Image Enhancement

As images become important data in daily life, image enhancement has become the focus of researchers. As early as the 1990s, many common contrast enhancement methods appeared, like improving image contrast by histogram equalization [33-35]. Subsequently, researchers improved the basic histogram method to enhance the image. In [36], the Contrast-Limited Adaptive Histogram Equalization was proposed. This method is the class of histogram stretching methods and is used to limit the contrast enhancement results of histogram equalization. The optimization technique, optimal contrast-tone mapping, for mapping the contrast tones of an image using a mathematical transfer function was proposed later [37].

After that, the field of image enhancement began to apply relatively complex techniques to improve the enhancement quality. The non-linear gamma function in enhancing the image contrast [37], stretching the histogram [38], or processing methods such as Brightness Preserving Bi-Histogram Equalization and Quantized Bi-Histogram Equalization [39] were proposed. These methods made a huge impact on improving the artifacts of histogram equalization. Afterward, for the problem of enhanced image denoising, the researchers also tried the methods of block matching and 3D filtering [40, 41], Singular Value Decomposition [42] as well as non-linear filters [43, 44], and achieved good results.

With the evolution of deep learning, it has gained world-renowned results in image recognition and other fields. Attempting to use deep learning methods to improve image enhancement techniques has become the focus of research in image enhancement in the last decade. First, the authors of [45] tried to use deep learning techniques to propose the concept of denoising autoencoders, which learn image features from noise to enhance image quality. In [46], CNN was applied to denoise images for the first time [47], and subsequently, methods of image inpainting [48] and deblurring [49] using CNN were also proposed.

Since then, image enhancement has introduced more sophisticated, advanced deep learning methods. The emergence of these advanced methods provides us with more ideas for solving image enhancement problems and inspires this paper's birth. For example, Lore proposed a method for enhancing low-light images using stacked sparse denoising autoencoders low-light net [50]. The low-light image enhancement method utilizing the Fully Connected Networks [51] opens the prelude to using deep learning networks with special advantages to improve low-light images. The state-of-the-art generative adversarial network mode averaging generative adversarial network approach [52], used for enhancing low-light images in a completely unsupervised manner, also shows the great potential of generative adversarial networks in the sphere of low-light image enhancement. The technique proposed in [53] showed good results with the enhancement of low-light images by using a very effective unsupervised generative adversarial network, dubbed EnlightenGAN, that can be trained

without low/normal-light image pairs yet proved to generalize very well on various real-world tests images. Another image enhancement algorithm without paired supervision is based on the Dark to Bright Generative Adversarial Network (D2BGAN) [54]. Using geometric and lighting consistency, along with a contextual loss criterion combined with multiscale color, texture, and edge discriminators, showed competitive results [55]. Enhanced traditional GANs with spectral normalization and advanced loss functions made the training stable and led to accurate results [55]. Experiments showed that the method proposed in [55] works very well qualitatively and quantitatively and alleviates saturation problems in bright areas, a typical problem of traditional low-light enhancement techniques. In paper [56], the proposed image enhancement method based on the enhanced network optimized generative adversarial network is put forward for solving the low-light image problem.

3. Proposed image enhancement algorithm

In this section, we introduce the composition principle of the Improved depthwise separable convolution generative adversarial networks, including how to refine the depthwise separable convolution module and the whole improved depthwise separable convolution generative adversarial network structure.

3.1. Improved depthwise separable convolution module

Fig. 5 shows the traditional convolution (a), the depthwise separable convolution in MobileNetV1 (b), and the contrast of the improved depthwise separable convolution structure in this paper (c). Compared with traditional convolution, depthwise convolution, and 1x1 convolution add the BN and ReLU activation layers. However, depthwise separable convolution in MobileNet is designed to deal with advanced vision problems, such as image classification; therefore, applying it to low-level vision tasks, such as image noise reduction and image enhancement, is impossible. Unlike high-level vision tasks, the input and output domains of low-light image enhancement tasks are both images; changes in network activation distribution, the so-called internal covariance shift, will not occur seriously during training. At the same time, batch standardization loses network flexibility, and image scale information also increases the loss of graphics processing units (GPU) in the training process. In this sense, batch normalization is no longer effective for low-light enhancement problems.

Therefore, this paper proposes an improved depthwise separable convolution module (IN-DepthwiseConv), as presented in Fig. 5c. First, instance normalization replaces batch normalization to preserve image scale information and deletes specific comparison information. This is performed to simplify the training procedure. Then, we delete the ReLU activation layer after the deep convolutional layer and keep the last ReLU activation layer to avoid losing information.

At the same time, inspired by MobileNetV2, to facilitate the componentization of the depthwise separable convolution module (Fig. 6a), this paper proposes an improved inverted residual depthwise separable convolution module (IN-Bottleneck) for the image enhancement problem, as shown in Fig. 6b.

By introducing instance normalization, the residual separable convolution module originally used to solve high-level vision problems can also be adapted to low-level vision problems, such as our target low-light image enhancement. Meanwhile, because the instance normalization does not contain trainable parameters, IN-Bottleneck does not increase the overall number of parameters or reduce the algorithm's efficiency.

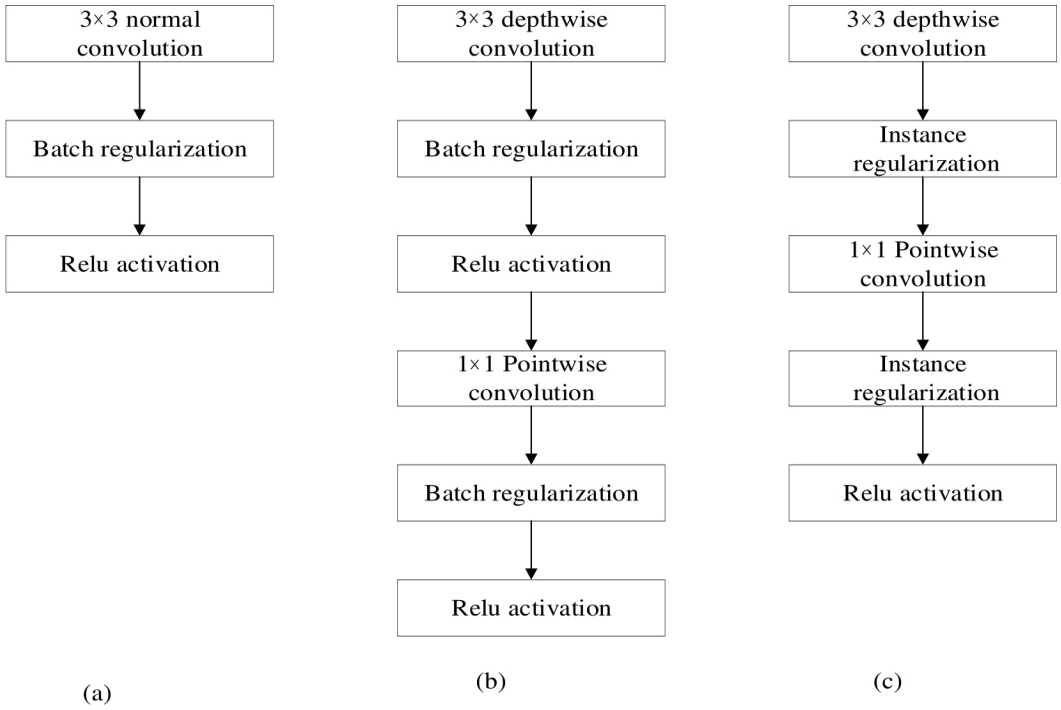

Fig. 5. Comparison of improved depthwise separable convolution structures: (a) traditional convolution layer, (b) depthwise separable convolution module, (c) improved depthwise separable convolution module.

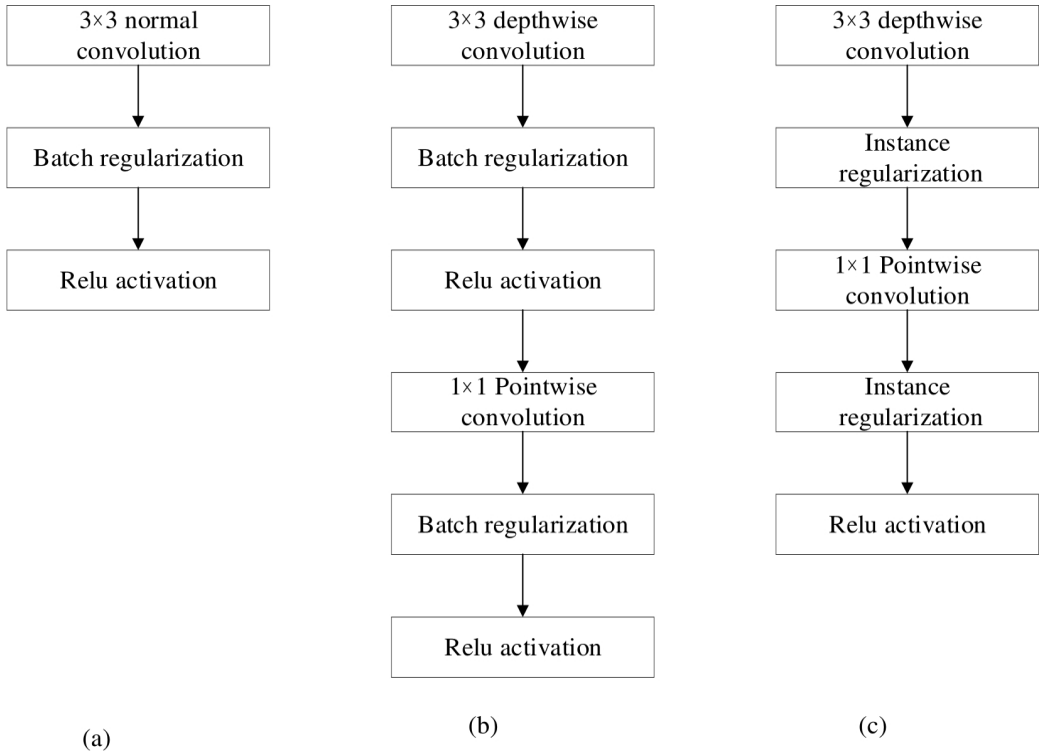

Fig. 6. Comparison of structures of the improved inverse residual depthwise separable convolution modules: (a) residual depwise separable convolution, (b) improved depthwise separable convolution module.

We implement the IN-DepthwiseConv module in the form of an algorithm 1:

ALGORITHM 1: IN-DepthwiseConv

Require: The low-light image
Ensure: Generate image features

```
Convolution module
    3×3 Depthwise convolution
    Instance normalization module
        Instance normalization
        PointWise module
            1×1 pointwise convolution
Instance normalization module
    Instance normalization
ReLu6 module
    ReLU6 Activation
Add up
Get the output (feature of image)
End
```

During the construction of our algorithm, IN-Bottleneck will also be implemented in the form of an algorithm. The specific process is as follows (algorithm 2):

ALGORITHM 2: IN-Bottleneck

Require: The low-light image
Ensure: Generate image features

```
Pointwise module
    1×1 pointwise convolution
    Instance normalization
    ReLU6 activation
DepthWise module
    3×3 Depthwise convolution
    Instance normalization
    ReLU6 activation
Linear module
    1×1 pointwise convolution
Instance normalization
Linear normalization
Add up
Get the output (feature of image)
```

3.2. Improved depthwise separable convolution generative adversarial networks

In paper [56], the method based on the enhanced network optimized generative adversarial network is put forward for solving the low-light image problem, and its advantages are proved; we name it LightAtten-GAN here. Based on its network structure, the improved depthwise separable convolution module is used to refine it further and get our image enhancement algorithm.

The LightAtten-GAN applies an improved attention mechanism, combined with the generative adversarial network algorithm, to comprehensively advance the enhancement effect on low-light images. However, the disadvantage is that the overall running speed is slow. So, we introduce its algorithm as our baseline algorithm and improve it by applying the two depthwise separable convolution mechanisms such as IN-DepthwiseConv and IN-Bottleneck.

The specific operation process of LightAtten-GAN is as follows (algorithm 3):

ALGORITHM 3: LightAtten-GAN

Require: The low-light image, the real image
Ensure: The enhanced image
Repeat
  Training generator network
    M low-light image pairs were randomly selected
    Generate the input of the fixed discriminant network with a length of *M*
    Calculate the total loss of the generator network
  Train the discriminator network
    Randomly input the initialized discriminant network with M stands for length
    M low-light image pairs were randomly selected
    Maximizes the overall loss of the discriminator network
Until
  Maximum number of iterations done
End

Then, we try to find the best way to combine them with the LightAtten-GAN. The generator and the discriminator need to be replaced. The *pros* and *cons* of the two depthwise separable convolution mechanisms must also be estimated. Therefore, we still need to build a comparison algorithm further to obtain the optimal combination. We constructed three different combinations of algorithms for comparative experiments, which we call DwDG, DwG, and DeeperDwG algorithms, respectively. The DwDG algorithm uses IN-DepthwiseConv and the IN-Bottleneck to replace the corresponding convolution layer in the generator and discriminator. The DwG algorithm only replaces the generator layer with the IN-DepthwiseConv module. The DeeperDwG algorithm replaced the generator with the deeper IN-Bottleneck module. We have compared the three algorithms in the experimental part to find the best algorithm.

Fig. 7 shows the lightweight residual unit in the generator, using an improved inverse residual depth separable convolution unit to replace the residual unit in a traditional generator (such as a traditional convolution network) to achieve the effect of a lightweight generator.

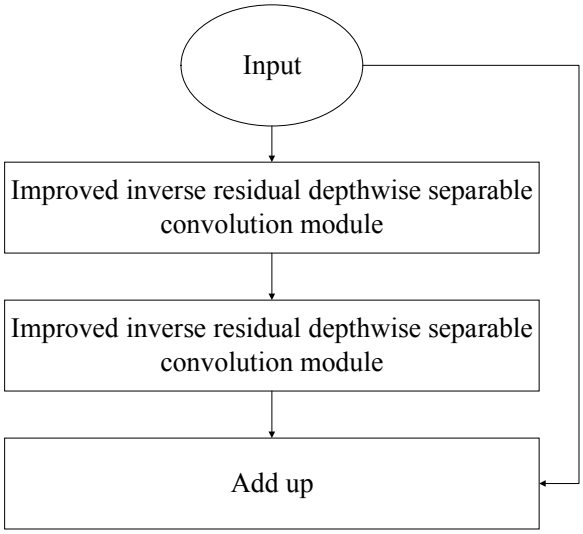

Fig. 7. The structure diagram of the Lightweight residual module.

Concerning parameters, the parameter amount of a residual unit of a traditional generator is:

$$Para_{Res} = (D_K \times D_K \times M) \times N \times 2 . \quad (2)$$

Among them, D_K, M, N represent the size of the convolution kernel, the number of channels of the input feature map, and the number of channels of the output feature map, respectively. The parameters of the lightweight residual unit in this chapter are as follows:

$$Para_{Depthwise0} = [M \times N(D_K \times D_K \times M) + M \times N] \times 2 . \quad (3)$$

It can be seen that the ratio of the parameters of the lightweight residual unit to the original unit is:

$$\frac{Para_{Depthwise0}}{Para_{Res}} = \frac{1}{N} + \frac{2}{D_K^2} . \quad (4)$$

Fig. 8 shows the network structure of the lightweight discriminator. In comparison with the traditional discriminator, the overall network is much lighter. Since the input size of the initial discriminator is (100, 100, 3) and the number of channels is small, we use the IN-Bottleneck with the same size of the deep convolution kernel and the original convolution kernel to replace the first convolution layer of the original discriminator. The purpose is to increase the number of channels before deep convolution and avoid data collapse problems. As the number of channels increases, we use an improved depthwise separable convolutional layer to replace the remaining convolutional layers. Thus, despite the lightweight network, the amount of $(1 \times 1 \times M) \times N$ parameters per layer can be further reduced.

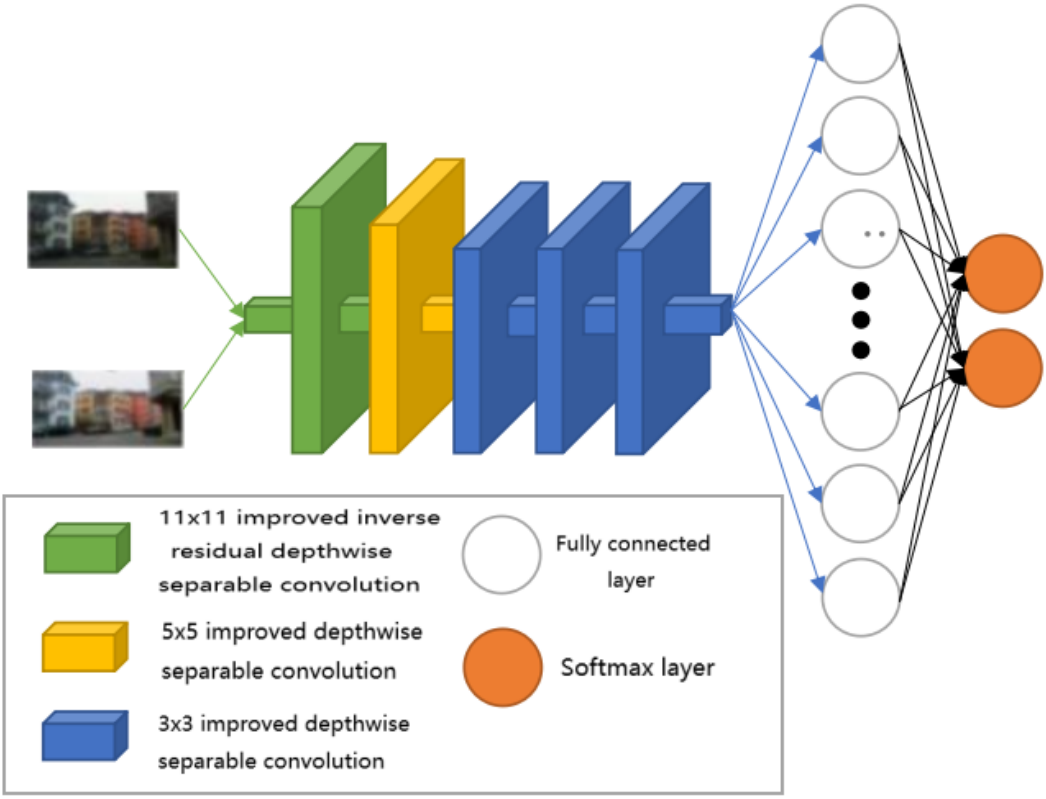

Fig. 8. The schematic diagram of the lightweight discriminator.

4. Performance evaluation

4.1. Experimental settings

A high-quality image data set with the resolution of DIVerse 2K (DIV2K) was chosen for the experiment images, including the training set of 800 pictures and the verification set of 100 pictures. To test the clear results of the algorithm for low-light image recognition accuracy, this experiment preprocessed DIV2K with low illumination and, specifically, utilized it to synthesize the low-light image data set.

The low-light images possess two important features: low-light and noise. In the pre-processing procedure, the adjustment of the image's low brightness relies on gamma correction and random parameters. The following is the formula:

$$I_L = \text{rand} \times (I_H)^{\gamma}, \tag{5}$$

where I_L - is the low-light image, rand - is the random number between (0,1), I_H - is the high-resolution image, γ - is the gamma coefficient, uniformly distributed between [1.1, 2].

Regarding the noise brought about by low-light, we also put the Gaussian noise with a uniform distribution of variance [0.01, 0.05] into the images treated with low-light. When the pre-processing ended, we prepared the training set of 30,744 images and the test set of 1080 images on the DIV2K data set. The image's size is 100×100, as presented in Figs. 9, 10.

Fig. 9. The original DIV2K image set.

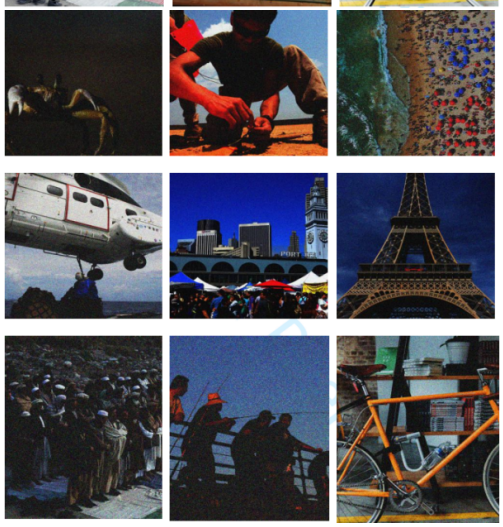

Fig. 10. The Div2k image set after low-light preprocessing.

The low-light paired data set (LOL) contains 500 pairs of low-light/normal-light images chosen by us, which is a dataset of images serving as low-light enhanced shots in real-life scenes. As a real image set, this data set can more truly reflect the effect of low-light image enhancement by the network after lightweight transformation.

We also used the digital photo enhancement dataset (DPED), which contains 22 K photos, including 4549 photos taken by Sony smartphone, 5727 photos taken by iPhone, and 6015 photos taken by Canon camera and BlackBerry. These photos were taken in automatic mode under various lighting conditions and weather in the daytime. All photos were collected using the camera's default settings.

We have carried out the experiment in this article through Tensorflow. The network proposed in this article can converge quickly and has been trained on NVIDIA GeForce GTX1080 for 20000 generations using a synthetic dataset. To prevent overfitting, we use flipping and rotation for data augmentation. We set the batch size to 32 and scaled the input image values to [0,1]. We use the 4th layer of the 5th convolution module in the VGG-19 network as the perceptual loss extraction layer. In the experiment, we used the Adam optimizer for training and employed a learning rate decay strategy, reducing the learning rate by 50% when the loss metric stopped improving. At the same time, to stabilize Gan training, we use spectral normalization and gradient penalty to constrain the discriminator.

Peak signal-to-noise ratio (*PSNR*) and structural similarity (*SSIM*) for quantitative analysis are used to assess the quality of the enhanced images. They are defined as follows: *PSNR* is an objective criterion used to evaluate images whose unit is dB. The calculation formula is as follows:

$$PSNR = 10\log_{10}\frac{(2^n-1)^2}{MSE}. \tag{6}$$

where n refers to the amount of bits of each pixel value, *MSE* represents the mean square error, whose calculation formula is as follows:

$$MSE = \frac{1}{w\,h}\sum_{i=0}^{w-1}\sum_{j=0}^{h-1}\| X(i,j)-Y(i,j)\|^2. \tag{7}$$

Among them, $X(i,j)$ and $Y(i,j)$ represent the source image and the target image, respectively, and w and h are the width and height of the image, respectively. If the *PSNR* value is larger, it represents less distortion and higher quality of the restored target image.

SSIM is a comprehensive image brightness, contrast, and structural difference in evaluating the similarity of two images' structures. The mathematical expression is as follows:

$$SSIM(X,Y) = 1(X,Y)^{\alpha}\, c(X,Y)^{\beta}\, s(X,Y)^{\gamma}. \tag{8}$$

In Eq. (8), X and Y refer to the source image and the target image, respectively, $1(X, Y)$, $c(X, Y)$, $s(X, Y)$ demonstrate the image brightness, contrast and structure difference, respectively. Their calculation formulae are as follows:

$$1(X,Y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2+\mu_y^2+C_1}, \tag{9}$$

$$c(X,Y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2+\sigma_y^2+C_2}, \tag{10}$$

$$s(X,Y)=\frac{2\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3}, \tag{11}$$

where μ_x,μ_y - show the average brightness of the source image *x* and the target image *y*, σ_x,σ_y - show the brightness standard deviation of the image *x*, *y*, and σ_{xy} - refers to the covariance between the images *x* and *y*. C_1,C_2,C_3 - are constants set to avoid the denominator of being 0, generally set $\alpha=\beta=\gamma=1$.

The *SSIM* value lies within the range of [0,1]. The larger the value, the better the restored image effect.

To evaluate the performance of lightweight, frames per second (*FPS*) and model calculation performance in giga floating point operations per second (*GFLOPS*) are calculated. For *FPS*, a larger value indicates faster inference speed. The calculation formula is as follows:

$$FPS=\frac{1}{Latency}, \tag{12}$$

where, *Latency* is the calculation time of the model, which is the average time required for an image to pass through the model calculation.

For computing model calculation performance in *GFLOPS*, we input the same batch of images into the model. We use the model weight calculation package profile in the Python framework to calculate the model's performance.

4.2. Experiments on synthetic datasets

First, we compare the baseline algorithm (LightAtten-GAN) with the three training algorithms we proposed, namely DWDG, DWG, and DeeperDWG, in the synthetic data set. We can see the quantitative results from Table 3 and the qualitative results from Figs. 11 and 12.

Table 3. Comparison of the experimental results obtained with LightAtten-GAN algorithm and proposed by us algorithms.

	LightAtten-GAN(BaseLine)	**DwDG**	**DwG**	**DeeperDwG**
PSNR	22.16	20.33	21.75	22.17
SSIM	0.999	0.979	0.993	0.999
Model Size (MB)	1.53	0.34	0.34	0.91

Because the DwG method only lightens the generator, the LightAtten GAN has decreased by 1.88% and 0.6% in PNSR and SSIM metrics, respectively, compared to before lightweight. However, the model size was reduced by 77.8%, indicating the effectiveness of our proposed lightweight method. The main reason is to ensure the discriminator's ability to distinguish between real and fake images, thereby enhancing the generator's low light enhancement capability.

We introduced multiple lightweight residual units in the generator to reduce the accuracy loss caused by lightweight and further improve image quality. The results showed that the DeeperDwG method remained the same as the LightAtten-GAN before lightweight in terms of PNSR and SSIM metrics but reduced the model size by 40.5%. This indicates that stacking multiple lightweight residual units increases the model complexity and brings

additional computational and memory overhead but also strengthens the model's ability to model low-light enhancement and improves the quality of generated images.

In Fig. 11, the red and blue boxes represent two partially enlarged images. It can be found that the three methods in this chapter can all improve the brightness of low-illumination images and denoise to a certain extent. But the DeeperDwG has the best effect, and the enhanced picture has proper brightness, balanced color, and clear details. For example, in the second row of pictures, compared with the real picture, the overall color difference between DwDG and DwG pictures is relatively large, and the blue box partially zoomed-in picture is blurry, losing detailed information. At the same time, the DeeperDwG has clear partial details and better overall visual effects. In the third row of pictures, compared to the real picture, the partially enlarged image after the DwDG and the DwG enhancement still has more noise, while the partially enlarged image of the DeeperDwG is more realistic. The petal patterns in the red box are more prominent, while the black gaps in the blue box are clearer and smoother. The overall effect even surpasses LightAtten GAN, approaching the real image.

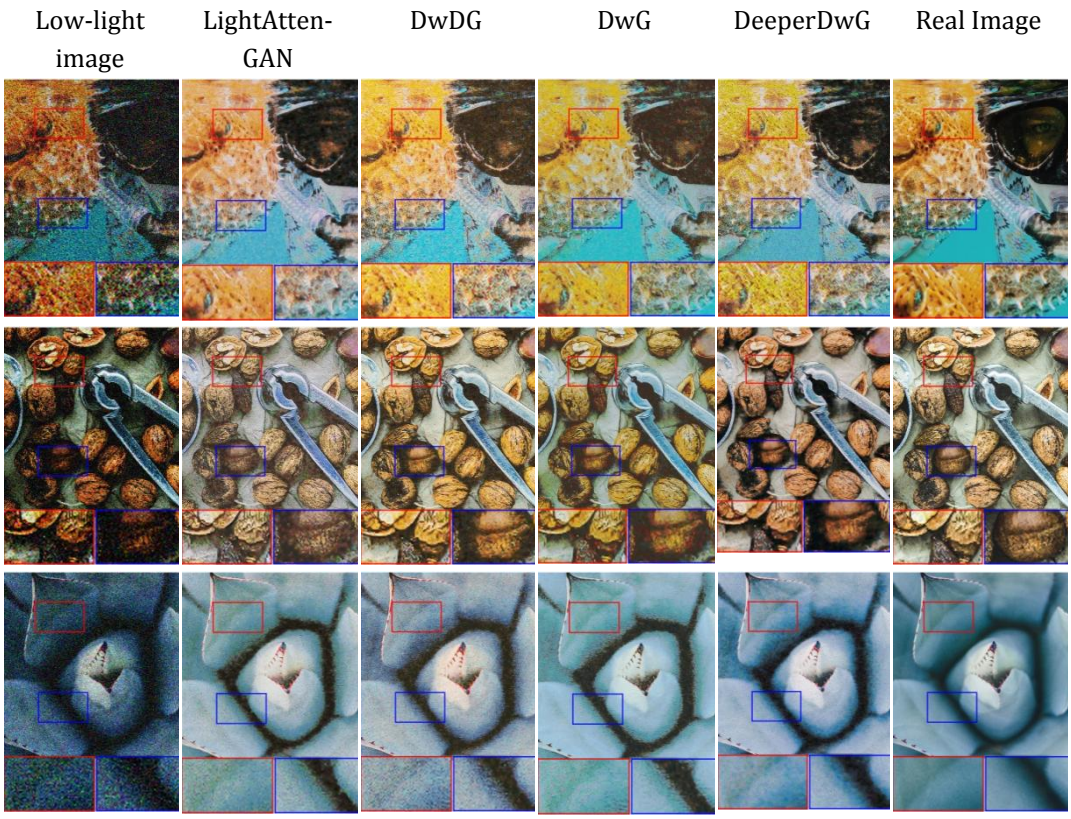

Fig. 11. Visual comparison over the synthetic dataset.

Fig. 12 is the brightness attention map of the experiment. The whiter area corresponds to the low-illuminance area of the original image, and the darker area corresponds to the high-brightness part of the original image. It is found that the brightness distribution of the brightness attention map predicted by the LightAtten-GAN and DeeperDwG is the same as the low-illuminance map. For example, in the low-illuminance original image in the first row, the

brightness of the oxygen mask is darker, and the brightness of the metal around the eye frame is obvious. Therefore, the oxygen mask part in the brightness attention map predicted by the DeeperDwG is brighter, and a black outline appears in the eye frame area. On the other hand, the oxygen mask part predicted by the DwDG method is darker overall, and the oxygen mask eye frame part predicted by the DwG method is not noticeably black. It shows that DeeperDwG is more effective than DwDG in predicting the brightness of the attention map.

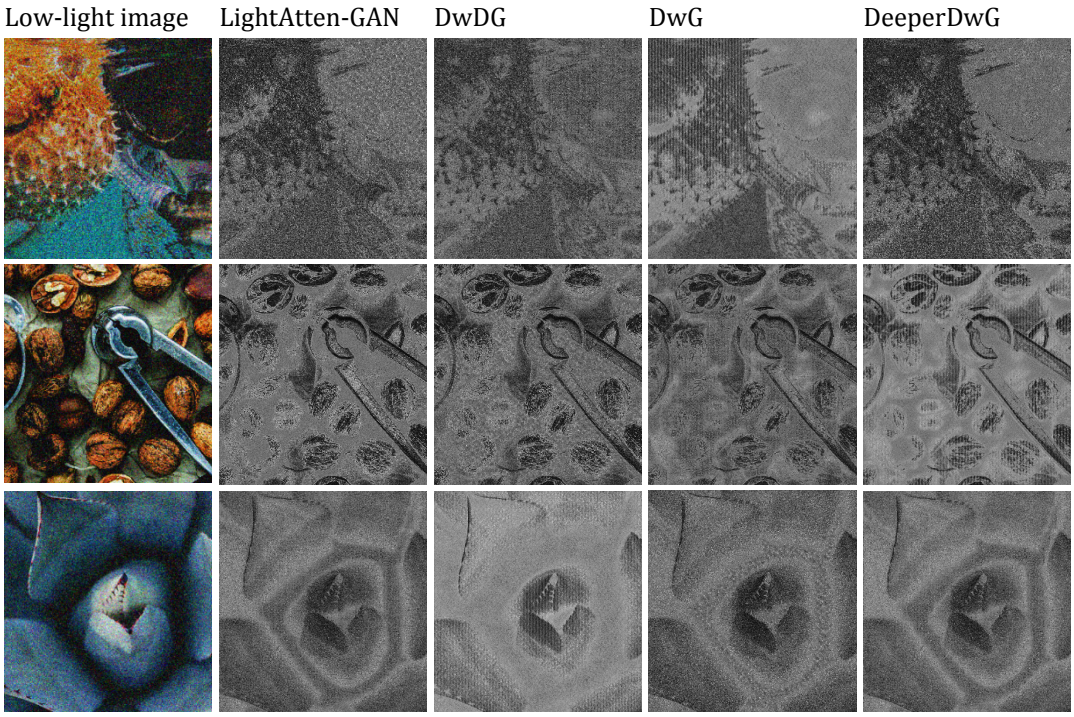

Fig. 12. Brightness attention contrast map on the synthetic data set.

4.3. Experiments on real datasets.

From the LOL dataset, the LightAtten-GAN algorithm is compared with the three proposed training algorithms, DwDG, DwG, and DeeperDwG algorithms. Table 5 presents the quantitative results, and Fig. 13 presents the qualitative results.

Table 4. Comparison of the experimental results obtained with LightAtten-GAN algorithm, and DwDG, DwG and DeeperDwG experimental results in LOL dataset.

	LightAtten-GAN(BaseLine)	**DwDG**	**DwG**	**DeeperDwG**
PSNR	20.99	19.85	20.12	20.91
SSIM	0.9986	0.9881	0.9901	0.9989
Model Size(MB)	2.2	0.41	1.2	1.6

Table 4 lists the PSNR, SSIM, algorithm sizes of the three methods, and the LightAtten-GAN on the LOL dataset. It should be noted that due to the relatively small LOL dataset, the algorithm size is much reduced compared to the Div2k dataset. At the same time, since this section only replaces the Div2k dataset with the LOL dataset, the DwDG, DwG, DeeperDwG network structure has not changed, so the algorithm size is the same as the performance on the Div2k dataset, so it is not listed.

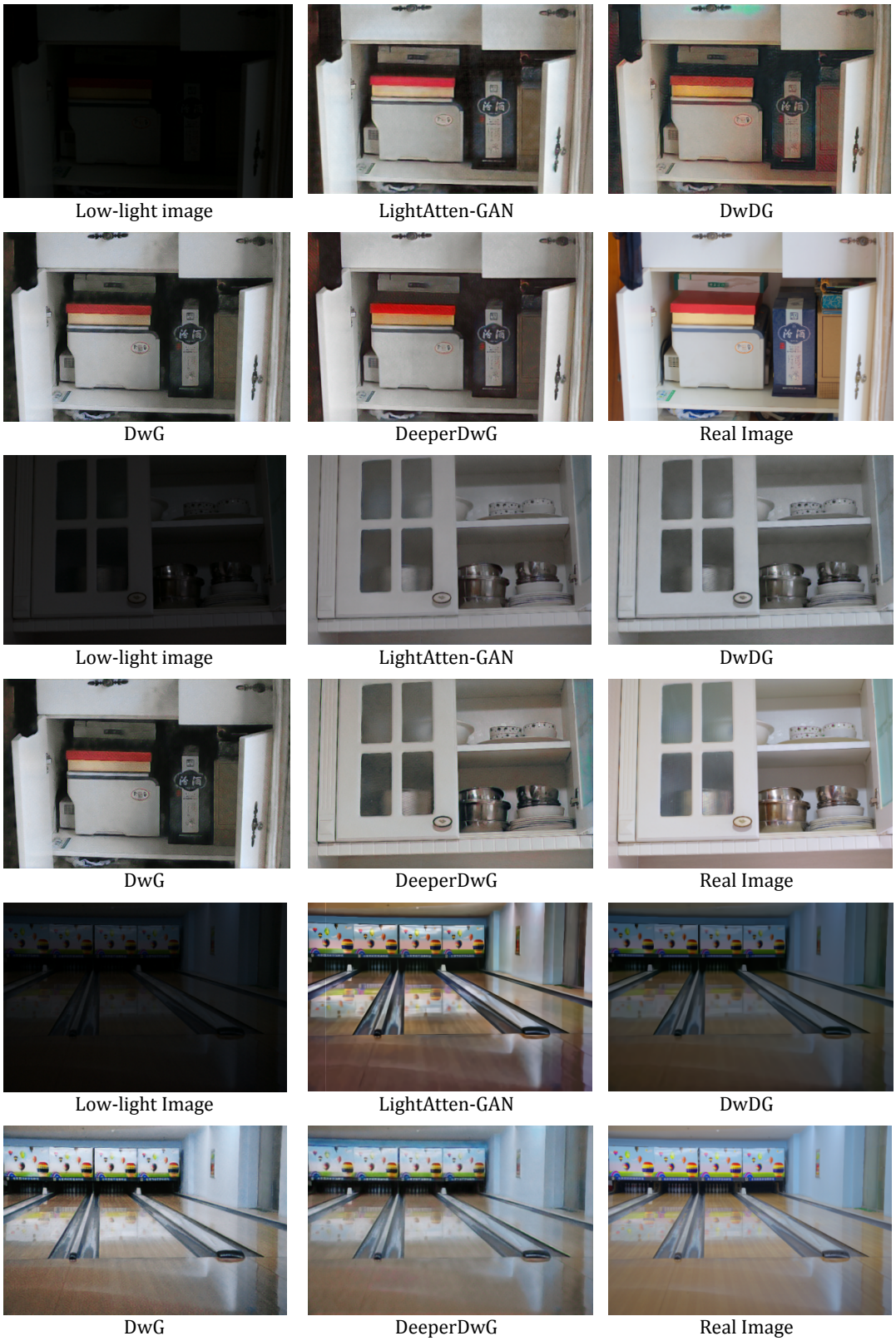

Fig. 13. Visual comparison of the DwDG, DwG, and DeeperDwG on the LOL dataset.

It can be found from Table 4 that the DeeperDwG method shows a slight decrease in PNSR metrics compared to LightAtten-GAN but a slight improvement in SSIM metrics. Additionally, it reduces model size by 40.5%, indicating the effectiveness of this method in accelerating computation and reducing the number of model parameters. Compared to DwG, the DwDG method has improved in both PNSR and SSIM metrics, indicating that the DeeperDwG method has a stronger ability to model low-light enhancement.

We compared our proposed method with the LightAtten-GAN on DPED. Table 5 presents the quantitative results, and Fig. 14 presents the qualitative results.

Table 5. Comparison results on DPED.

	LightAtten-GAN	**DwDG**	**DwG**	**DeeperDwG**
PSNR	20.46	19.58	20.1`	20.35
SSIM	0.9156	0.8572	0.9045	0.9116
Model Size(MB)	5.52	0.86	2.57	3.34

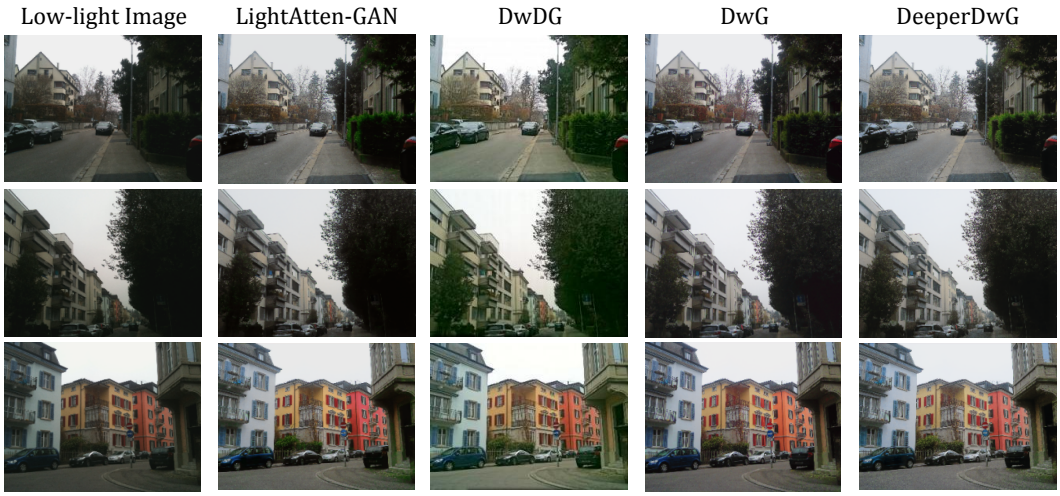

Fig. 14. Visual comparison on the DPED.

DwDG, DwG, and DeeperDwG are all effective in reducing model size, with DwDG showing the largest reduction of nearly 84.4%. However, their PSNR and SSIM also showed significant decreases, with reductions of 4.3% and 6.37%, respectively, consistent with the no-free lunch theorem. Furthermore, although the DeeperDwG method showed a slight decrease in PSNR and SSIM compared to LightAtten GAN, it was generally consistent overall.

Fig. 14 shows the visual effect of the LightAtten-GAN and our three methods on the DPED. From the figure, we can see that the use of the DwDG has a large color difference, and the picture is generally green, especially when there are more green plants and trees in the picture; the reason may be that the DPED is collected outdoors, containing more natural and living scenes, while the DwDG method has weakened the ability of the discriminator to distinguish true and false images to some extent. As a result, the generated image is greenish as a whole. In addition, we found that the image generated by the DeeperDwG method has a better visual effect and color saturation than that generated by the DwG method. For example, in the first row of the image, the street vegetation in the image generated by DeeperDwG is more saturated and natural than that generated by the DwG method, and the overall contrast of the street is obvious.

Finally, we compared our methods with the LightAtten-GAN on DPED. Our methods have a slight decrease in PSNR and SSIM when compared with the LightAtten-GAN, but it is basically the same in general. At the same time, the algorithm size has been considerably reduced, which shows that it has significantly reduced the number of parameters. It can also be seen from the visual effects in Fig. 12 that there is almost no difference between the actual DeeperDwG and other images, which shows that DeeperDwG compensates for the accuracy loss caused by a lightweight network and ensures the quality of the enhanced images.

Table 6. Comparison results on DPED.

	LightAtten-GAN	DwDG	DwG	DeeperDwG
PSNR	20.46	19.58	20.10	20.35
SSIM	0.9156	0.8572	0.9045	0.9116
Model Size(MB)	5.52	0.86	2.57	0.91
FPS	29	62	53	51
GLOP	23.4	27.6	27.1	26.8

5. Conclusions and future research

This article proposes an improved depthwise separable convolutional generation network designed to fix the problem of excessive calculation at low illumination image enhancement and excessive algorithm. In many real-world applications, such as execution on computing-constrained platforms, the low-light enhancement algorithm based on CNN has computational complexity and memory problems. Therefore, the network introduces depthwise separable convolution and improves it to reduce algorithm parameters. It is also suitable for low-light image enhancement tasks. Inspired by the MobileNet series of networks, we proposed the IN-DepthwiseConv module and IN-Bottleneck. Based on these designed modules, three algorithms were proposed: the DwDG, DwG, and DeeperDwG. The DwDG algorithm uses the improved depth separable convolution and improved inverse residual depth separable convolution modules to substitute the corresponding convolution layers in the generator and discriminator. The DwG algorithm only replaces the generator's convolutional layer. The DeeperDwG algorithm uses the generator's deeper inverted residual depth separable convolution module. Finally, the method proposed in this paper confirms the algorithm's performance in the synthetic and real data sets (LOL) and is compared with the other methods. In the final part, it is shown the network we proposed can guarantee the low-light image enhancement effect and reduce the number of algorithm parameters. This reduces the computational complexity. In the experimental part, we propose a practical case of applying the DeepDwG to engineering. This proves that our proposed algorithm fully satisfies the application of multimedia computing in urban scenarios.

Although our method enhances low-light images effectively, it has the limitation that it is difficult to deal with various lighting modes and styles. In the future, we will further explore new methods to improve low-light image enhancement for trustworthy multimedia computing urban scene applications.

Funding. Not applicable.

Disclosures. The authors declare no conflict of interest.

References

1. Rybář, J., Hučko, B., Ďuriš, S., Pavlásek, P., Chytil, M., Furdová, A., & Veselý, P. (2020). Factors affecting measurements of IOP using non-contact eye tonometer. *Strojnícky časopis - Journal of Mechanical Engineering, 70*(2), 133–140.

2. Jun, S., Kochan, O., & Kochan, R. (2016). Thermocouples with built-in self-testing. *International Journal of Thermophysics, 37*(4), 37.
3. Song, W., Beshley, M., Przystupa, K., Beshley, H., Kochan, O., Pryslupskyi, A., Pieniak, D., & Su, J. (2020). A software deep packet inspection system for network traffic analysis and anomaly detection. *Sensors, 20(6), 1*637.
4. Krolczyk, G. M., & Legutko, S. (2014). Experimental analysis by measurement of surface roughness variations in turning process of duplex stainless steel. *Metrology and Measurement Systems, 21*(4), 759-770.
5. Su, J., Beshley, M., Przystupa, K., Kochan, O., Rusyn, B., Stanisławski, R., Yaremko, O., Majka, M., Beshley, H., Demydov, I., & Kahalo, I. (2022). 5G multi-tier radio access network planning based on Voronoi diagram. *Measurement, 192*, 110814.
6. Abdullah-Al-Wadud, M., Kabir, M. H., Dewan, M. A. A., & Chae, O. (2007). A dynamic histogram equalization for image contrast enhancement. *IEEE Transactions on Consumer Electronics, 53*(2), 593–600.
7. Kwan, C., Larkin, J., & Ayhan, B. (2020). Demosaicing of CFA 3.0 with applications to low lighting images. *Sensors, 20*(12), 3423.
8. Kwan, C., Larkin, J., & Budavari, B. (2020, April). Demosaicing images in low lighting environments. In *Signal Processing, Sensor/Information Fusion, and Target Recognition XXIX* (Vol. 11423, pp. 267-284). SPIE.
9. Kwan, C., & Larkin, J. (2019). Demosaicing of Bayer and CFA 2.0 patterns for low lighting images. *Electronics, 8*(12), 1444.
10. Land, E. H., & McCann, J. J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America, 61*(1), 1–11.
11. Jobson, D. J., Rahman, Z., & Woodell, G. A. (1997). Properties and performance of a center/surround retinex. *IEEE Transactions on Image Processing, 6*(3), 451–462.
12. Jobson, D. J., Rahman, Z. U., & Woodell, G. A. (1997). A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image Processing*, *6*(7), 965-976.
13. Fu, Q., Jung, C., & Xu, K. (2018). Retinex-based perceptual contrast enhancement in images using luminance adaptation. *IEEE Access*, *6*, 61277-61286.
14. Yuan, L., & Sun, J. (2012). Automatic exposure correction of consumer photographs. In *Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part IV 12* (pp. 771-785). Springer Berlin Heidelberg.
15. He, K., Sun, J., & Tang, X. (2011). Single image haze removal using dark channel prior. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(12), 2341–2353. https://doi.org/10.1109/TPAMI.2010.168
16. Li, C., Fan, T., Ma, X., Zhang, Z., Wu, H., & Chen, L. (2017, June). An improved image defogging method based on dark channel prior. In *2017 2nd international conference on image, vision and computing (ICIVC)* (pp. 414-417). IEEE.
17. Lore, K. G., Akintayo, A., & Sarkar, S. (2017). LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognition*, *61*, 650-662.
18. Shen, L., Yue, Z., Feng, F., Chen, Q., Liu, S., & Ma, J. (2017). MSR-net: Low-light image enhancement using deep convolutional network. *arXiv*. https://arxiv.org/abs/1711.02488
19. Cai, J., Gu, S., & Zhang, L. (2018). Learning a deep single image contrast enhancer from multi-exposure images. *IEEE Transactions on Image Processing*, *27*(4), 2049-2062.
20. Wei, C., Wang, W., Yang, W., & Liu, J. (2018). Deep Retinex decomposition for low-light enhancement. arXiv. https://arxiv.org/abs/1808.04560
21. Lv, F., Lu, F., Wu, J., & Lim, C. (2018, September). MBLLEN: Low-light image/video enhancement using cnns. In *BMVC* (Vol. 220, No. 1, p. 4).
22. Chen, C., Chen, Q., Xu, J., & Koltun, V. (2018). Learning to see in the dark. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3291-3300).
23. Zhang, Y., Zhou, C., Chang, F., & Kot, A. C. (2019). Multi-resolution attention convolutional neural network for crowd counting. *Neurocomputing*, *329*, 144-152.
24. Zhou, Z., Feng, Z., Liu, J., & Hao, S. (2020). Single-image low-light enhancement via generating and fusing multiple sources. *Neural Computing and Applications*, *32*(11), 6455-6465.
25. Fang, M. T., Chen, Z. J., Przystupa, K., Li, T., Majka, M., & Kochan, O. (2021). Examination of abnormal behavior detection based on improved YOLOv3. *Electronics*, *10*(2), 197.
26. Li, H., He, X., Tao, D., Tang, Y., & Wang, R. (2018). Joint medical image fusion, denoising and enhancement via discriminative low-rank sparse dictionaries learning. *Pattern Recognition*, *79*, 130-146.
27. Yan, L., Li, K., Gao, R., Wang, C., & Xiong, N. (2022). An intelligent weighted object detector for feature extraction to enrich global image information. *Applied Sciences*, *12*(15), 7825.

28. Ulyanov, D., Vedaldi, A., & Lempitsky, V. (2016). Instance normalization: The missing ingredient for fast stylization. arXiv. https://arxiv.org/abs/1607.08022
29. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv. https://arxiv.org/abs/1704.04861
30. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4510-4520).
31. Zhang, X., Zhou, X., Lin, M., & Sun, J. (2018). Shufflenet: An extremely efficient convolutional neural network for mobile devices. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 6848-6856).
32. Ma, N., Zhang, X., Zheng, H. T., & Sun, J. (2018). Shufflenet v2: Practical guidelines for efficient cnn architecture design. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 116-131).
33. Trahanias, P. E., & Venetsanopoulos, A. N. (1992, January). Color image enhancement through 3-D histogram equalization. In *11th IAPR International Conference on Pattern Recognition. Vol. III. Conference C: Image, Speech and Signal Analysis,* (Vol. 1, pp. 545-548). IEEE Computer Society.
34. Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., Romeny, B.H., Zimmerman, J.B. & Zuiderveld, K. (1987). Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing*, *39*(3), 355-368.
35. Cheng, H. D., & Shi, X. J. (2004). A simple and effective histogram equalization approach to image enhancement. *Digital Signal Processing*, *14*(2), 158-170.
36. Pisano, E. D., Zong, S., Hemminger, B. M., DeLuca, M., Johnston, R. E., Muller, K., Braeuning, M. P., & Pizer, S. M. (1998). Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms. *Journal of Digital imaging*, *11*, 193-200.
37. Gonzalez, R. C., & Woods, R. E. (2001). *Digital image processing* (2nd ed.). Prentice Hall. Upper Saddle River, NJ.
38. Krutsch, R., & Tenorio, D. (2011). Histogram equalization. *Freescale Semiconductor, Document Number AN4318, Application Note*, *30*.
39. Kaur, M., Kaur, J., & Kaur, J. (2011). Survey of contrast enhancement techniques based on histogram equalization. *International Journal of Advanced Computer Science and Applications*, *2*(7).
40. Dabov, K., Foi, A., Katkovnik, V., & Egiazarian, K. (2009, April). BM3D image denoising with shape-adaptive principal component analysis. In *SPARS'09-Signal Processing with Adaptive Sparse Structured Representations*.
41. Dabov, K., Foi, A., Katkovnik, V., & Egiazarian, K. (2008, March). Image restoration by sparse 3D transform-domain collaborative filtering. In *Image Processing: Algorithms and Systems VI* (Vol. 6812, pp. 62-73). SPIE.
42. Elad, M., & Aharon, M. (2006). Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, *15*(12), 3736-3745.
43. Chen, T., Ma, K. K., & Chen, L. H. (1999). Tri-state median filter for image denoising. *IEEE Transactions on Image Processing*, *8*(12), 1834-1838.
44. Cheng, H. D., & Shi, X. J. (2004). A simple and effective histogram equalization approach to image enhancement. *Digital Signal Processing, 14*(2), 158–170.
45. Vincent, P., Larochelle, H., Bengio, Y., & Manzagol, P. A. (2008, July). Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th International Conference on Machine Learning* (pp. 1096-1103).
46. Jain, V., & Seung, S. (2008). Natural image denoising with convolutional networks. *Advances in Neural Information Processing Systems*, *21*.
47. Xie, J., Xu, L., & Chen, E. (2012). Image denoising and inpainting with deep neural networks. *Advances in neural information processing systems*, *25*.
48. Schuler, C. J., Hirsch, M., Harmeling, S., & Schölkopf, B. (2014). Learning to deblur. *arXiv preprint arXiv:1406.7444*. https://arxiv.org/abs/1406.7444
49. Agostinelli, F., Anderson, M. R., & Lee, H. (2013). Adaptive multi-column deep neural networks with application to robust image denoising. *Advances in Neural Information Processing Systems*, *26*.
50. Lore, K. G., Akintayo, A., & Sarkar, S. (2017). LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognition*, *61*, 650-662.
51. Chen, C., Chen, Q., Xu, J., & Koltun, V. (2018). Learning to see in the dark. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3291-3300).
52. Wang, R., Jiang, B., Yang, C., Li, Q., & Zhang, B. (2022). MAGAN: Unsupervised low-light image enhancement guided by mixed-attention. *Big Data Mining and Analytics*, *5*(2), 110-119.
53. Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., Yang, J., Zhou, P., & Wang, Z. (2021). Enlightengan: Deep light enhancement without paired supervision. *IEEE transactions on image processing*, *30*, 2340-2349.

54. Bhattacharya, J., Modi, S., Gregorat, L., & Ramponi, G. (2022). D2bgan: A dark to bright image conversion model for quality enhancement and analysis tasks without paired supervision. *IEEE Access*, *10*, 57942-57961.
55. Kim, G., Kwon, D., & Kwon, J. (2019, September). Low-lightgan: Low-light enhancement via advanced generative adversarial network with task-driven training. In *2019 IEEE International Conference on Image Processing (ICIP)* (pp. 2811-2815). IEEE.
56. Yan, L., Fu, J., Wang, C., Ye, Z., Chen, H., & Ling, H. (2021). Enhanced network optimized generative adversarial network for image enhancement. *Multimedia Tools and Applications*, *80*, 14363-14381.

Y. Qiu, H. Wang, T. Demkiv, O. Kochan, L. Yan. (2025). Low-Light Image Enhancement Based on Depthwise Separable Convolution**.** *Ukrainian Journal of Physical Optics*, *26*(1), 01040 – 01063. doi: 10.3116/16091833/Ukr.J.Phys.Opt.2025.01040.

Анотація*. З моменту появи алгоритмів глибокого навчання, алгоритми на основі згорткових нейронних мереж (CNN) суттєво просунулися в покращенні зображення при слабкому освітленні. Однак вони все ще стикаються з серйозною проблемою: алгоритм покращення зображення при низькій освітленості на основі CNN має надмірну обчислювальну складність і потребує достатньої пам'яті. Хоча точність алгоритму покращується, ефективність обчислень знижується. У цій статті представлено легку мережу для низького освітлення та запропоновано покращення зображення. Для ознайомлення подані основи використовуваної технології. Базуючись на принципі MobileNetV2, ми використали генеративні змагальні мережі з покращеними механізмами уваги, як базовий алгоритм. Потім були побудовані три порівняльні алгоритми для експериментів. Результати експерименту підтверджують, що запропонована мережа потребує менше параметрів алгоритму, гарантуючи ефект покращення зображення в умовах слабкого освітлення.*

Ключові слова*: машинне навчання, покращення зображення, змагальні мережі, роздільна глибинна згортка*