

Received March 15, 2021, accepted April 5, 2021, date of publication April 9, 2021, date of current version April 19, 2021. Digital Object Identifier 10.1109/ACCESS.2021.3072331

RetinexGAN:Unsupervised Low-Light Enhancement With Two-Layer Convolutional Decomposition Networks

TIAN MA[®]¹, (Member, IEEE), MING GUO[®]¹, ZHENHUA YU[®]¹, (Member, IEEE), YANPING CHEN[®]², XINCHENG REN³, (Member, IEEE), RUNTAO XI[®]¹, YUANCHENG LI[®]¹, AND XINLEI ZHOU[®]¹

¹College of Computer Science and Technology, Xi'an University of Science and Technology, Xi'an 710054, China
²Shaanxi Key Laboratory of Network Data Analysis and Intelligent Processing, Xi'an University of Posts and Telecommunications, Xi'an 710121, China
³Shaanxi Key Laboratory of Intelligent Processing for Big Energy Data, Yan'an University, Yan'an 716000, China

Corresponding author: Ming Guo (gmingyun@foxmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61834005, in part by the Enterprise Joint Fund Project of Shaanxi Natural Science Basic Research Plan under Grant 2019JLM-11-2, and in part by the Shaanxi Key Laboratory of Intelligent Processing for Big Energy Data under Grant IPBED11.

ABSTRACT In the field of image enhancement, using deep learning methods to enhance low-light images is currently mainstream. However, the methods often have complex network structures of a large number of parameters, and their training often uses paired data-sets, which are difficult to obtain in actual practice. To solve these problems, this paper proposes a simple generative adversarial network and Retinex model, dubbed RetinexGAN, that is completely trained using unpaired data-sets. It contains a decomposition network and two discriminator networks. To reduce the parameters of the network, only two convolution layers are used in the decomposition network. We show more challenging testing data where some parts of the image are underexposed and others are normal light. Both quantitative and visual results show that RetinexGAN is largely superior to state-of-the-art methods.

INDEX TERMS Low-light image enhancement, generative adversarial networks, two-layer convolutional networks.

I. INTRODUCTION

When shooting an image, the quality of the image will be greatly reduced by factors such as insufficient light, photography equipment performance and photographing skills, which will lead to low visibility, low contrast and noise. Human visual perception is challenged not only by these problems but also by the accuracy of some computer vision algorithms. To improve image quality, there are currently many hardware devices and software technologies that are dedicated to making the details hidden in low-light images clearly presented. These methods have broader application prospects. For example, in photography, low-light image enhancement technology can help users shoot high-quality and more attractive images in low-light environments; for workers working in low-light environments, low-light enhancement technology can make their working environment clearer on the monitoring equipment to ensure their safety; and in artificial

The associate editor coordinating the review of this manuscript and approving it for publication was Ze Ji¹⁰.

intelligence, especially in the field of computer vision, it can obviously improve target detection and recognition accuracy.

Early low-light enhancement methods were mainly based on a histogram or Retinex theory. The method based on histogram equalization only considers pixel-by-pixel mapping, and ignores that pixels will be affected by surrounding pixels. The Retinex method first predicts an illumination map and performs image reconstruction through this illumination map. However, this method is a complex mathematical problem. It is difficult to predict an illumination map under extremely dark conditions with unavoidable noise.

Recent research mainly adopts the deep learning method to learn color, contrast, brightness and saturation to produce more expressive results. It mainly relies on synthetic or captured damaged/clean image pairs for training. However, these methods still have limitations:

(1) The synthesized paired data-sets are often not realistic enough, and it is difficult to produce good results when the trained model is applied to real low-light images. Paired images are extremely difficult to obtain in reality,



FIGURE 1. Some representative results in RetinexGAN, which can not only enhance the low-light areas in the image but also maintain the normal-light areas.

and the generalization ability of their trained models is often poor.

(2) A good result requires a complex neural network, which will produce a large number of model parameters during the training process. However, in some mobile devices, to ensure a high device storage performance, a smaller model file is often required.

In this paper, we propose a new learning-based low-light enhancement method. The network we use is simple in structure, fast in training. And it is trained on unpaired data-sets. FIGURE 1 shows the results of using our method to enhance a low-light image. Our contributions include the following:

(1) The decomposition network only needs two convolution layers to decompose the eigenimage. The low-light image is decomposed into a reflection map and an illumination map.

(2) Unpaired images are used in the network for low-light enhancement. This training strategy eliminates the dependence on paired training data so that the model has better generalization and can use more types of images from different fields for training.

(3) The network structure is simple, and the lightweight one-way GAN network model greatly reduces the training complexity and training time, which can be trained with no more than 100 epochs. In our work, the size of the trained model is 15KB-150KB. In addition, we can still adjust the parameters to achieve different enhancement results in tests. Therefore, for different data-sets, we do not need to retrain the model which can be applied to mobile devices with small memory and flexible applications.

FIGURE 2 proves the effectiveness of the two-layer convolution. Generated Gaussian noise is inputted into the VGG-19 pretrained model to extract features. As shown in FIGURE 2, the first to fifth rows use one-layer, two-layer, three-layer, four-layer, and seventeen-layer convolutions. The first to sixth columns are 5 to 30 iterations. Only the first two convolution layers retain color information.



FIGURE 2. Features extracted by different convolutional layers.

II. RELATED WORK

Research on image enhancement has a long history. Wang *et al.* [1] compared the low-light image enhancement methods in recent decades and divided them into seven categories. According to the research hotspots in recent years, we introduce four methods: histogram equalization, Retinex, methods based on deep learning, and methods based on GAN.

A. TRADITIONAL METHODS

In the past few decades, many scholars have endeavored to use histogram equalization and methods based on Retinex. The histogram equalization methods attempt to map the entire histogram into a simple mathematical distribution. The difference between different histogram equalization methods is the use of different additional priorities and constraints. There are some newer methods, such as DOTHE (dominant orientation-based texture histogram equalization) [2], EDSHE (entropy-based dynamic subhistogram equalization) [3] and UMHE (unsharp masking with histogram equalization) [4]. These methods mainly enhance the contrast of the image but insufficiently enhance the illumination, which leads to under/over enhancement or color loss.

Methods based on the Retinex theory assume that the image is composed of reflection maps and illumination maps. Some typical methods, such as MSR (multi-scale Retinex) [5] and SSR (single-scale Retinex) [6] use illumination maps for low-light image enhancement. AMSR(adaptive multiscale Retinex) [7] is proposed as a weighting strategy based on SSR. NPE (naturalness preserved enhancement) [8] balances the enhancement level and the naturalness of an image to avoid excessive enhancement. Fu *et al.* [9] proposes a model named SRIE (simultaneous reflectance and illumination estimation), which can accurately maintain the estimated reflection coefficient and suppress noise to a



FIGURE 3. Network structure. *x* is the low-light image. $x_{max/min/mean}$ is the minimum, maximum, and mean value of the RGB channels in *x*. x^R is a reflection map. x^I is an illumination map. *y* is a normal image. y_{max}^3 is the maximum value of the RGB channels in *y*. \otimes denotes pixel-by-pixel multiplication. \oplus denotes pixel-by-pixel addition. The decomposition network uses two convolution layers. The first convolution layer has 6 input channels and 64 output channels. The second convolution layer has 64 input channels and 6 output channels.

certain extent. LIME (low-light image enhancement) [10], a structure-aware smoothing model, is designed to estimate the illumination map, and uses the reflection map as the final enhancement result. BIMEF (bio-inspired multi-exposure fusion) [11] proposes a double exposure fusion algorithm, and Ying *et al.* [12] uses the camera response model for further enhancement. Wang *et al.* [13] proposes an adaptive low-light image enhancement method. They converts the image into HSV space and then estimated the illumination map of the image. The method based on the Retinex theory can effectively enhance low-light images. However, the key of these Retinex-based methods is to estimate the illumination map. Mathematically, estimating the illumination map is an ill-posed problem. Moreover, the parameters of this method are handcrafted, relying on careful parameter adjustment.

B. METHODS BASED ON DEEP LEARNING

With the maturity of deep learning technology, it has been increasingly applied to the field of image enhancement, which has led to new developments in low-light image enhancement algorithms. LLNet (low-light net) [14] is the first network to apply deep learning to image enhancement in a true sense. It enhances and denoises low-light images by learning a sparse adaptive encoder. HDRNet (high definition resolution net) [15] predicts the coefficients of the local affine model in the bilateral space to obtain low-light image enhancement. LLCNN (low-light convolutional neural network) [16], [17] relys on some traditional methods during training and is not an end-to-end solution. MSRNet (multiscale Retinex net) [18] uses different Gaussian convolution kernels to learn dark to bright mapping directly. The MBLLEN (multibranch low-light enhancement network) [19] uses a multibranch low-light enhancement

network structure to extract features of different levels, enhance them through multiple subnetworks, and generate output images through multibranch fusion. RetinexNet [20] is the first method to combine Retinex theory with CNN, which achieves the purpose of low-illumination map enhancement by estimating and adjusting the illumination map. Similarly, KinD (kin-dling the darkness) [21] also applies Retinex theory to estimate the illumination map by adding a restoration network for noise removal and then enhance the low-light image. Wang et al. [22] introduces intermediate lighting in their network (DeepUPE) and correlats the input with the expected enhancement result, which enhanced the network's ability to learn complex photographic adjustments from paired data-sets modified by experts. Ren et al. [23] proposes a deep hybrid network to improve image detail and edge and uses perceptual and adversarial losses to improve image quality. Zhang et al. [24] is inspired by the theory of information entropy and the retinal model and proposes a Retinex model based on maximum entropy, which uses a simple network to separate illumination and reflection to achieve self-supervised learning. Huang et al. [25] uses the attention mechanism to predict the illumination map, and uses the Retinex model to estimate the initial enhanced image. This method uses the image after histogram equalization as the reference image so that the enhanced image has chromatic aberration. However, most of the above methods use paired data-sets for training. Guo et al. [26] proposes a lightweight deep network named DCE-Net (deep curve estimation net), which is used to estimate pixels and depict high order curves to adjust the dynamic range of a given image. The attraction of DCE-Net is that it does not require any paired or unpaired data during training but achieves better enhancement results. However, generated image is whitish.

C. METHOD BASED ON GAN

Currently, GANs have achieved excellent results in imageto-image conversion work [27], [28]. In the low-light image enhancement field, Chen et al. [29] regards the image enhancement problem as an image-to-image conversion problem, and proposes a two-way GAN network that uses unpaired images enhance low-light images. Different from two-way GAN, EnlightenGAN [30] proposes a lightweight one-way GAN. It uses a UNet with self-feature retention loss and a self-regular attention mechanism and two global-local discriminators to achieve unsupervised low-light image enhancement. Yang et al. [31] proposes the DRBN (deep recursive band network) method, which uses paired low-light/normal-light images to restore and enhance the linear band representation of normal-light images and then uses other unpaired data driven by perceptual quality-driven adversarial learning. A linear transformation can be learned to reconstruct a given band to obtain an improved band representation. However, some areas of the image enhanced by this method will become blurred. Moreover, after testing, its enhancement results on real-world data-sets [32] are not ideal.

However, many current GAN network structures are more complicated. The GAN network has disadvantages such as troublesome tuning and poor stability. A more complex network structure will inevitably increase the difficulty of network training. Our goal is to find a simpler network to achieve the purpose of low-light image enhancement. Unlike EnlightenGAN, our method only uses two convolution layers to decompose the eigenimage. Faster training will be achieved and more pleasing images will be produced.

III. APPROACH

A. RETINEX WITH CORRECTION COEFFICIENT

In the original Retinex theory, an image S is regarded as the product of illumination map I and reflection map R. The formula is:

$$S_{input} = I * R, \tag{1}$$

where S_{input} is the input low-light image, * represents pixelby-pixel multiplication, the reflection map R is a constant determined by the nature of the object, and the illumination map I is affected by the external light source, which can be removed by removing the light influence or lighting component I is corrected to achieve the purpose of image enhancement.

Most current methods use the estimated reflection map R as an enhanced (good exposure) image [10]. However, we found in experiments that using only the reflection map will cause the loss of image details, color, and excessive enhancement. To solve this problem, we made the following improvements to the original Retinex formula:

$$S_{output} = \beta * (1 - I) * R + \alpha * S_{input}, \qquad (2)$$

this formula is used in testing instead of training. It is simplified yet effective. The result is shown in FIGURE 4, where



FIGURE 4. Using *input*, R and (1 - I) as the parameters of FORMULA 2, we can get the enhanced image as shown in the fifth line of the figure.

I and *R* are the illumination map and the reflection map, respectively. They are estimated using the decomposition network (the specific structure is introduced in the next section). We normalize *I* and *R* to [0, 1] and use (1 - I) to suppress the lighter areas and lighten the darker areas. S_{input} is the original low-light image, and S_{output} is the enhanced image. β and α are the image brightness correction coefficients. The value range of β is [1, 8]. As the value increases, the brightness of the low-light area increases. The value range of α is [1, 3]. As the value increases.

It is important to point out that Retinex theory assumes that 1) the visual appearance can be decomposed by reflection maps and illumination maps, 2) reflectivity is the intrinsic property of the object, and 3) the illumination map can be varied(depending on actual conditions). Therefore, unlike the illumination map and reflection map of methods [21], [22], our reflection map is closer to the normal image, and our illumination map is closer to the grayscale image of the normal image. This situation is more similar to [20].

B. NETWORK STRUCTURE

FIGURE 3 shows our network structure. The left area is used for training. The right area is used for testing.

1) DECOMPOSITION NETWORK

Zeiler and Fergus [33] visualized the features extracted by the convolutional layers. Layer1 and layer2 learn the underlying features, such as edge, color, and brightness. Layer3 begins to learn high-level texture features. Layer4 learns some distinguishing features, and layer5 learns features with key distinctions. The features learned by the CNN are gradually captured in a higher level, as shown in FIGURE 2. Based on this,



FIGURE 5. Model structure of discriminator.

we design a decomposition network with a two-layer convolutional neural network. We only use layer1 and layer2 to extract the color and brightness of the image without changing the texture and contour of the original image.

Lv and Lu [34] believes that the quality of information in the illumination map is lower. To solve the problem of less information in the illumination map, unlike other single-channel illumination maps based on Retinex theory, our illumination map uses 3 channels. We extract the minimum, maximum, and mean values of the RGB channels from the low-light image before training, and concatenate them to maintain more image details in the illumination map.

2) DISCRIMINATOR NETWORK

After experimenting, we found that the use of ordinary discriminators may overenhance the brighter areas or underenhance the darker areas in the image. To ensure that our network can adaptively enhance different regions in the image, we adopt the discriminator structure of PatchGAN, namely, the Markov discriminator. The discriminator structure is shown in FIGURE 5.

Currently, the Markov discriminator is used in GAN networks such as pix2pix [28] and CycleGAN [27]. It has a certain high resolution and high detail retention for super high resolution and clear pictures in style transfer. The discriminator maps the input to an N * N patch (matrix) X, which is actually the feature map output by the convolutional layer. The value of X_{ij} represents the probability that each patch is a true sample, and the average value of X_{ij} is the final output of the discriminator. Each output in the output matrix represents a receptive field in the original image, corresponding to a patch of the original image.

We use two PatchGAN discriminators as the illumination map discriminator and the reflection map discriminator. The two discriminators have the same structure, but the parameters are different. They are used to judge the authenticity of the generated illumination map and reflection map. This not only ensures that the content (semantics) similarity is

VOLUME 9, 2021

maintained between the generated image and the original image but also ensures that the illumination map and reflection map of the low-light image can learn the brightness information of the normal-light images.

C. LOSS FUNCTION

The loss function of the network in this paper is only composed of adversarial loss. The adversarial loss is used to make $x^{I} \sim \mathbb{P}_{\text{real-patches}}$, which means that the distribution of the low illumination map is approximate to that of the normal image. Additionally, we need to make $x^{R} \sim \mathbb{P}_{\text{real-patches}}$, which means that the distribution of the reflection map is close to that of the normal image.

For the local discriminator, sixteen patches are randomly cut from the output image and the real image. Here, we use the original least square GAN (LSGAN) [35] as the antagonistic loss. LSGAN, as its name suggests, uses the objective function as a mean square error. The discriminator no longer maximizes the cost function of the original GAN but to minimizes it. The objective is to learn both the forward mapping $G^I : \{x^I\} \rightarrow y_{max}^3$ and $G^R : \{x^R\} \rightarrow y$. In this paper, we code the generated samples and the real samples as 0 and 1. Using square error as the objective function, the goal of the illumination map discriminator is as follows:

$$\min \mathcal{L}\left(D^{\mathrm{I}}\right) = \frac{1}{2} \mathbb{E}_{x^{I} \sim \mathbb{P}_{\mathrm{fake-patches}}} \left[\left(D^{\mathrm{I}}\left(x^{I}\right) - 0\right)^{2} \right] \\ + \frac{1}{2} \mathbb{E}_{y_{\mathrm{max}}^{3} \sim \mathbb{P}_{\mathrm{real-patches}}} \left[\left(D^{\mathrm{I}}\left(y_{\mathrm{max}}^{3}\right) - 1\right)^{2} \right],$$
(3)

$$\min \mathcal{L}\left(G^{\mathrm{I}}\right) = \frac{1}{2} \mathbb{E}_{x^{I} \sim \mathbb{P}_{\mathrm{fake-patches}}} \left[\left(D^{\mathrm{I}}\left(x^{I}\right) - 1\right)^{2} \right].$$
(4)

Similarly, the goal of the reflection map discriminator is:

$$\min \mathcal{L}\left(D^{\mathsf{R}}\right) = \frac{1}{2} \mathbb{E}_{x^{\mathsf{R}} \sim \mathbb{P}_{\mathsf{fake-patches}}} \left[\left(D^{\mathsf{R}}\left(x^{\mathsf{R}}\right) - 0\right)^{2} \right] + \frac{1}{2} \mathbb{E}_{y \sim \mathbb{P}_{\mathsf{real-patches}}} \left[\left(D^{\mathsf{R}}(y) - 1\right)^{2} \right], \quad (5)$$



FIGURE 6. Low-light image as input. The third and fifth columns are parts of the enlarged image. (a) input, (b) SRIE [9], (c) LIME [10], (d) RetinexNet [20], (e) SIEN [24], (f) EnlightenGAN [30], (g) DCE-Net [26], (h) DRBN [31], (i) ours, $\beta = 6.2$, $\alpha = 1.3$. Please zoom in to see the details.

$$\min \mathcal{L}\left(G^{\mathsf{R}}\right) = \frac{1}{2} \mathbb{E}_{x^{\mathsf{R}} \sim \mathbb{P}_{\mathsf{fake-patches}}} \left[\left(D^{\mathsf{I}}\left(x^{\mathsf{R}}\right) - 1\right)^{2} \right].$$
(6)

where x^{I} is generated illumination map and x^{R} is generated reflection map, respectively. y_{max}^{3} is the maximum RGB value

IEEE Access



FIGURE 7. Normally illuminated image as input. (i) are input images. (a) SRIE [9], (b) LIME [10], (c) RetinexNet [20], (d) SIEN [24], (e) EnlightenGAN [30], (f) DCE-Net [26], (g) DRBN [31], (h) ours, $\beta = 3$, $\alpha = 1.1$. Please zoom in to see the details.

of the normal illumination map cascaded into three channels. D^{I} and D^{R} are the discriminators of illumination map and reflection map.

The total loss function of our decomposition network is as follows:

$$\mathcal{L} = \mathcal{L}\left(G^{\mathrm{I}}\right) + \mathcal{L}\left(G^{\mathrm{R}}\right). \tag{7}$$



 $\beta=1$, $\alpha=1.5$

 $\beta=3, \alpha=1.5$

 β =5, α =1.5

 β =3.5, α =2.5 | β =6, α =1.3

FIGURE 8. Results obtained with different β and α . Please zoom in to see the details.

 $\mathcal{L}(G^{I})$ and $\mathcal{L}(G^{R})$ are the adversarial losses of illumination and reflection.

IV. EXPERIMENT

A. DATASET AND EXPERIMENTAL DETAILS

We use unpaired low/normal-light images for training, and the data-set is provided by EnlightenGAN [30]. These images are converted to PNG format and adjusted to 600*400 pixels. We select 900 low-light images and 900 normal-light images from this data-set as the training set. To expand the data-set, the images are randomly cropped to 400*400 and flipped during training. For the testing set, we compare previous works (LIME [10], LOL [20], SRIE [9], DeepUPE [22], DICM [36], and EnlightenGAN [30]) using standard images.

Our batch size is set to 1. We use the Adam optimizer. The learning rate of the decomposition network is set to 1e-4. The learning rate of the two discriminator networks is set to 4e-4. The network needs to train 100 epochs on a single Nvidia 1080Ti GPU, and it takes no more than 1 hour.

One of the shortcomings of GANs is that the training is unstable. Since the network structure we used is a variant of training. After many repeated experiments, we found that the model achieves better results when iterating for approximately 50 epochs. We suggest that readers should save the model at intervals during the training process and check the generated reflection map when replicating the code in this paper. When the color, brightness, and texture of the reflection map meet your expectations, the model parameters are optimal at this time. Of course, readers can also refer to other articles that optimize GAN stability [37]–[39] to make training more stable.

the GAN network, this problem is inevitable in our network

B. EXPERIMENTAL COMPARISON

1) COMPARISON WITH TYPICAL EXPERIMENTS

We compare our method with the following state-of-the-art low-light image enhancement methods: SRIE [9], LIME [10], RetinexNet [20], SIEN [24], EnlightenGAN [30], DCE-Net [26], DRBN [31]. To ensure fairness in comparison, we use the source code and pretrained models provided by the authors of these methods and use the parameter settings recommended by the authors. That data-set in FIGURE 6 is from

 TABLE 1. Quantitative comparison with other low-light enhancement methods. NIQE ie expected smaller values. PSNR and SSIM are expected higher values.

	NIQE	PSNR	SSIM
SRIE	9.25	16.85	0.74
LIME	10.10	17.44	0.82
RetinexNet	10.97	13.01	0.71
SIEN	14.34	9.64	0.65
EnlightenGAN	9.21	16.77	0.80
DCE-Net	8.97	18.03	0.75
DRBN	14.03	18.45	0.81
Ours	8.15	19.33	0.82

TABLE 2. Pretrained model size and runtime comparisions (in second).

	Model size	Running speed	Platform
SRIE	-	7.14s	MATLAB(CPU)
LIME	-	0.19s	MATLAB(CPU)
RetinexNet	9267KB	1.09s	TensorFlow(CPU)
SIEN	2479KB	0.81s	TensorFlow(CPU)
EnlightenGAN	33774KB	0.36s	PyTorch(GPU)
DCE-Net	313KB	0.19s	PyTorch(GPU)
DRBN	20638KB	3.60s	PyTorch(GPU)
Ours	137KB	0.13s	PyTorch(GPU)

DeepUPE [22]. We can see in the FIGURE 6 that SRIE [9] enhances the input image while its brightness enhancement effect is limited compared to other methods. LIME [10] overenhances image areas with light sources or areas with higher pixel values, resulting in loss of surrounding details. After using RetinexNet [20] for enhancement, the texture of the image is very unnatural. SIEN [24] overenhances and distorts the color. In EnlightenGAN [30], although the attention map is used in its work to eliminate image artifacts, there are still some artifacts in some test images, and the generated effect is yellowish. DCE-Net [26] is whitish. DRBN [31] loses some details of the image.

We also carried out a quantitative comparison. Unpaired images (from DeepUPE [22]) and paired images (Enlighten-GAN [30]) are used for testing. For paired images, we use PSNR and SSIM to evaluate image quality. For unpaired images, we use NIQE to evaluate image quality. The average values are shown in TABLE 1.

We tested the running speed of each method. As shown in TABLE 2, our method is the fastest. For the traditional method, we use a CPU (Intel Core i7-7700HQ CPU) for testing. Due to the limitations of the hardware environment, we also use a CPU to test RetinexNet and SIEN. The other methods are tested on an Nvidia GTX1050 Ti GPU.

The current methods were proposed to enhance low-light images, but they fail to maintain enhancement performance when we input images with normal light. As shown in FIGURE 7, (i) are the input normal-light images, and (a)-(h) are the enhanced results outputted by different methods. Obviously, we have a better performance than other approaches, e.g., the enhanced images in line (h) are closer to the images in line (i) than other lines.



FIGURE 9. The influence of β and α on NIQE, PSNR and SSIM.

2) ABLATION STUDY

Different people may like images with different brightness. However, most current methods based on deep learning can only obtain one kind of brightness image after the operation of low-light image enhancement. By simply adjusting the β and α , we can meet the requirements of most people for low-light image enhancement.

In the test, the different β and α values in FORMULA 2 will affect the test results. FIGURE 8 shows the different brightness enhancement effects caused by different β and α in the test. β is used to adjust the enhancement of dark



FIGURE 10. Face detection results before and after low-light image enhancement. Note that there is almost no brightness in the fourth low-light image. Please zoom in to see the details. $\beta = 8$, $\alpha = 1.3$.

areas in the image. As β increases, dark areas in the image are illuminated. This value can guarantee the texture of the image. However, if the β value is too large, it will cause the image details to be lost. α is used to control the overall enhancement of the image. As α increases, the overall brightness of the image will increase. This value can guarantee the details of the image. However, if the α value is too large, part of the image (especially at the light source) will be overexposed. In addition, we tested the effect of β and α on image quality. As shown in FIGURE 9, as the value of β increases, the value of NIQE will decrease, but the values of PSNR and SSIM will also decrease. When α is 2 and β is 1, PSNR and SSIM achieve optimal values. However, in reality, images with good visual effects may not have large PSNR and SSIM values [40].

C. REAL-WORLD IMAGE ENHANCEMENT

To verify the generalization of the model, the FD-LOL (face detection in low-light conditions) [32] data-set is used in the test. This data-set has real low-light images taken at night, and is used for face detection in low-light scenes. We used DSFD (dual shot face detector) [41] method to detect the face of the original low-light image and the image enhanced by RetinexGAN. FIGURE 10 shows before image enhancement, face detection missed detections and false detections.

After image enhancement, face detection accuracy is significantly improved.

V. CONCLUSION

In this paper, we propose a deep learning model based on generative adversarial networks and Retinex to enhance low-light images with unpaired data-sets. It uses a decomposition network with only two convolution layers to decompose a low-light image into an illumination map and a reflection map. Then, Retinex with correction coefficients is used to reconstruct the illumination map and the reflection map into normal images. Our experiments proved the superiority of this method. In future work, we will focus on adaptive and real-time low light image enhancement, and apply this model to multi-agent systems [42], cyberphysical systems [43], and automated manufacturing systems [44].

REFERENCES

- W. Wang, X. Wu, X. Yuan, and Z. Gao, "An experiment-based review of low-light image enhancement methods," *IEEE Access*, vol. 8, pp. 87884–87917, 2020.
- [2] K. Singh, D. K. Vishwakarma, G. S. Walia, and R. Kapoor, "Contrast enhancement via texture region based histogram equalization," *J. Mod. Opt.*, vol. 63, no. 15, pp. 1444–1450, Aug. 2016.
- [3] A. S. Parihar and O. P. Verma, "Contrast enhancement using entropy-based dynamic sub-histogram equalisation," *IET Image Process.*, vol. 10, no. 11, pp. 799–808, Nov. 2016.
- [4] S. Kansal, S. Purwar, and R. K. Tripathi, "Image contrast enhancement using unsharp masking and histogram equalization," *Multimedia Tools Appl.*, vol. 77, no. 20, pp. 26919–26938, Oct. 2018.
- [5] D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Trans. Image Process.*, vol. 6, no. 7, pp. 965–976, Jul. 1997.
- [6] D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Trans. Image Process.*, vol. 6, no. 3, pp. 451–462, Mar. 1997.
- [7] C.-H. Lee, J.-L. Shih, C.-C. Lien, and C.-C. Han, "Adaptive multiscale retinex for image contrast enhancement," in *Proc. Int. Conf. Signal-Image Technol. Internet-Based Syst.*, Dec. 2013, pp. 43–50.
 [8] S. Wang, J. Zheng, H.-M. Hu, and B. Li, "Naturalness preserved enhance-
- [8] S. Wang, J. Zheng, H.-M. Hu, and B. Li, "Naturalness preserved enhancement algorithm for non-uniform illumination images," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3538–3548, Sep. 2013.
- [9] X. Fu, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding, "A weighted variational model for simultaneous reflectance and illumination estimation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2782–2790.
- [10] X. Guo, Y. Li, and H. Ling, "LIME: Low-light image enhancement via illumination map estimation," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 982–993, Feb. 2017.
- [11] Z. Ying, G. Li, and W. Gao, "A bio-inspired multi-exposure fusion framework for low-light image enhancement," 2017, arXiv:1711.00591. [Online]. Available: http://arxiv.org/abs/1711.00591
- [12] Z. Ying, G. Li, Y. Ren, R. Wang, and W. Wang, "A new low-light image enhancement algorithm using camera response model," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2017, pp. 3015–3022.
 [13] W. Wang, Z. Chen, X. Yuan, and X. Wu, "Adaptive image enhance-
- [13] W. Wang, Z. Chen, X. Yuan, and X. Wu, "Adaptive image enhancement method for correcting low-illumination images," *Inf. Sci.*, vol. 496, pp. 25–41, Sep. 2019.
- [14] K. G. Lore, A. Akintayo, and S. Sarkar, "LLNet: A deep autoencoder approach to natural low-light image enhancement," *Pattern Recognit.*, vol. 61, pp. 650–662, Jan. 2017.
- [15] M. Gharbi, J. Chen, J. T. Barron, S. W. Hasinoff, and F. Durand, "Deep bilateral learning for real-time image enhancement," *ACM Trans. Graph.*, vol. 36, no. 4, pp. 1–12, Jul. 2017.
- [16] L. Tao, C. Zhu, G. Xiang, Y. Li, H. Jia, and X. Xie, "LLCNN: A convolutional neural network for low-light image enhancement," in *Proc. IEEE Vis. Commun. Image Process. (VCIP)*, Dec. 2017, pp. 1–4.
- [17] L. Tao, C. Zhu, J. Song, T. Lu, H. Jia, and X. Xie, "Low-light image enhancement using CNN and bright channel prior," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 3215–3219.

- [18] L. Shen, Z. Yue, F. Feng, Q. Chen, S. Liu, and J. Ma, "MSR-Net: Low-light image enhancement using deep convolutional network," 2017, arXiv:1711.02488. [Online]. Available: http://arxiv.org/abs/1711.02488
- [19] F. Lv, F. Lu, J. Wu, and C. Lim, "Mbllen: Low-light image/video enhancement using CNNs," in *Proc. BMVC*, 2018, p. 220.
- [20] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," 2018, arXiv:1808.04560. [Online]. Available: http://arxiv.org/abs/1808.04560
- [21] Y. Zhang, J. Zhang, and X. Guo, "Kindling the darkness: A practical low-light image enhancer," in *Proc. 27th ACM Int. Conf. Multimedia*, Oct. 2019, pp. 1632–1640.
- [22] R. Wang, Q. Zhang, C.-W. Fu, X. Shen, W.-S. Zheng, and J. Jia, "Underexposed photo enhancement using deep illumination estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 6849–6857.
- [23] W. Ren, S. Liu, L. Ma, Q. Xu, X. Xu, X. Cao, J. Du, and M.-H. Yang, "Low-light image enhancement via a deep hybrid network," *IEEE Trans. Image Process.*, vol. 28, no. 9, pp. 4364–4375, Sep. 2019.
- [24] Y. Zhang, X. Di, B. Zhang, and C. Wang, "Self-supervised image enhancement network: Training with low light images only," 2020, arXiv:2002.11300. [Online]. Available: http://arxiv.org/abs/2002.11300
- [25] W. Huang, Y. Zhu, and R. Huang, "Low light image enhancement network with attention mechanism and retinex model," *IEEE Access*, vol. 8, pp. 74306–74314, 2020.
- [26] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, and R. Cong, "Zeroreference deep curve estimation for low-light image enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1780–1789.
- [27] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2223–2232.
- [28] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1125–1134.
- [29] Y.-S. Chen, Y.-C. Wang, M.-H. Kao, and Y.-Y. Chuang, "Deep photo enhancer: Unpaired learning for image enhancement from photographs with GANs," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6306–6314.
- [30] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, "EnlightenGAN: Deep light enhancement without paired supervision," *IEEE Trans. Image Process.*, vol. 30, pp. 2340–2349, 2021.
- [31] W. Yang, S. Wang, Y. Fang, Y. Wang, and J. Liu, "From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 3063–3072.
- [32] J. Liang, J. Wang, Y. Quan, T. Chen, J. Liu, H. Ling, and Y. Xu, "Recurrent exposure generation for low-light face detection," 2020, arXiv:2007.10963. [Online]. Available: http://arxiv.org/abs/2007.10963
- [33] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2014, pp. 818–833.
- [34] F. Lv, Y. Li, and F. Lu, "Attention guided low-light image enhancement with a large scale low-light simulation dataset," 2019, arXiv:1908.00682. [Online]. Available: http://arxiv.org/abs/1908.00682
- [35] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, "Least squares generative adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2794–2802.
- [36] C. Lee, C. Lee, and C.-S. Kim, "Contrast enhancement based on layered difference representation," in *Proc. 19th IEEE Int. Conf. Image Process.*, Sep. 2012, pp. 965–968.
- [37] L. Mescheder, S. Nowozin, and A. Geiger, "The numerics of GANs," in Proc. Adv. Neural Inf. Process. Syst., 2017, pp. 1825–1835.
- [38] V. Nagarajan and J. Z. Kolter, "Gradient descent GAN optimization is locally stable," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5585–5595.
- [39] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," 2017, arXiv:1701.07875. [Online]. Available: http://arxiv.org/abs/1701.07875
- [40] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4681–4690.
- [41] J. Li, Y. Wang, C. Wang, Y. Tai, J. Qian, J. Yang, C. Wang, J. Li, and F. Huang, "DSFD: Dual shot face detector," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 5060–5069.

- [42] X. B. Li, Z. H. Yu, Z. W. Li, and N. Q. Wu, "Group consensus via pinning control for a class of heterogeneous multi-agent systems with input constraints," *Inf. Sci.*, vol. 542, pp. 247–262, 2021, doi: 10.1016/j.ins.2020.05.085.
- [43] S. A. Zhou, Z. H. Yu, E. S. A. Nasr, H. A. Mahmoud, E. M. Awwad, and N. Q. Wu, "Homomorphic encryption of supervisory control systems using automata," *IEEE Access*, vol. 8, pp. 147185–147198, 2020.
- [44] X. Zan, Z. P. Wu, C. Guo, and Z. H. Yu, "A Pareto-based genetic algorithm for multi-objective scheduling of automated manufacturing systems," *Adv. Mech. Eng.*, vol. 12, no. 1, pp. 1–15, 2020.



TIAN MA (Member, IEEE) was born in Henan, China, in 1982. He received the B.S. degree in measurement and control technology and instrument, the M.S. degree in software engineering, and the Ph.D. degree in information and communication engineering from Northwestern Polytechnical University, Xi'an, China, in 2003, 2006, and 2011, respectively. Since 2014, he has been an Associated Professor with the College of Computer Science and Technology, Xi'an University of

Science and Technology, China. He is currently the author of more than 20 articles and more than ten software copyrights. His research interests include image processing, computer graphics, and 3D simulation and visualization. He has served as the Organizing Committee Co-Chair for the 14th International Conference on Verification and Evaluation of Computer and Communication Systems (VECoS 2020) and a common Reviewer for the IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC).



MING GUO was born in Shandong, China, in 1996. He received the bachelor's degree in computer science and technology from the Xi'an University of Science and Technology, China, in 2019. He is currently pursuing the master's degree with the College of Computer Science and Technology, Xi'an University of Science and Technology. His research interest includes computer vision.



ZHENHUA YU (Member, IEEE) received the B.S. and M.S. degrees from Xidian University, Xi'an, China, in 1999 and 2003, respectively, and the Ph.D. degree from Xi'an jiaotong University, Xi'an, in 2006. He is currently a Professor with the College of Computer Science and Technology, Institute of Systems Security and Control, Xi'an University of Science and Technology, Xi'an. He has authored more than 20 technical papers for conferences and journals and holds two

invention patents. His research interests include cyber-physical systems, machine learning, and system security.



YANPING CHEN received the B.S. degree in computer software from the Xi'an University of Posts and Telecommunications, Xi'an, China, in 2001, and the Ph.D. degree in computer science and technologies from Xi'an Jiaotong University, Xi'an, in 2007.

From 2012 to 2013, she was a Visiting Research Faculty with Iowa State University, Ames, USA. She is currently a Full Professor and the Assistant Dean of the School of Computer Science,

Xi'an University of Posts and Telecommunications, Xi'an. She has authored over 70 articles. Her research interests include service computing, data representation, and data analysis.



XINCHENG REN (Member, IEEE) was born in 1967. He received the B.S. degree in physics and theoretical physics from Yanan University, in 1990, Yan'an, China, the M.S. degree in physics and theoretical physics from Northwest University, Xi'an, China, in 2000, and the Ph.D. degree in radio physics from Xidian University, Xi'an, in 2008.

He is currently a Professor with the School of Physics and Electronic Information, Yanan Uni-

versity. He has authored over 100 articles. His research interests include propagation and scattering of electromagnetic (optical) waves in complex systems and random medium, computational electromagnetics, and theory and technology of wireless communication.



YUANCHENG LI was born in Henan, China, in 1981. He received the B.S. degree in computer science and technology from the Xi'an University of Posts and Telecommunications, China, in 2004, and the M.Sc. degree in computer software theory and the Ph.D. degree in computer system structure from Xi'an Jiaotong University, China, in 2007 and 2012, respectively.

Since 2012, he has been a Lecturer with the College of Computer Science and Technology, Xi'an

University of Science and Technology, China. His research interests include parallel computing and compiler optimization. He serves as a Reviewer for *Concurrency and Computation: Practice and Experience* and the *Journal of the Chinese Institute of Engineers*.



RUNTAO XI received the bachelor's degree in software engineering from Langfang Normal University, in 2019. He is currently pursuing the master's degree with the Xi'an University of Science and Technology. His research interests include computer vision and reinforcement learning, especially adaptive image enhancement in reinforcement learning.



XINLEI ZHOU was born in Xi'an, Shaanxi, China, in 1997. He received the bachelor's degree in engineering from Anhui Polytechnic University, in 2020. He is currently pursuing the master's degree in engineering from the Xi 'an University of Science and Technology. His research interests include computer vision and data visualization.

...