IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 30 May 2024, accepted 8 June 2024, date of publication 18 June 2024, date of current version 25 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3416130

RESEARCH ARTICLE

Non-Uniform Low-Light Face Image Enhancement Based on Dark Channel Prior and Image Uniform Posterior

BO-YU ZHANG, QING-CHUN ZHANG, AND WEN-YING ZHANG
School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou 450001, China
Corresponding author: Wen-Ying Zhang (iewyzhang@zzu.edu.cn)

ABSTRACT Non-uniform low-light images, characterized by complex lighting conditions, pose a significant challenge in image restoration and enhancement, particularly under backlit scenarios where the high-brightness background obscures the foreground information. Existing methods struggle with the diverse illumination factors across different scenes. To overcome this, we introduce a novel deep reconstruction network designed specifically for enhancing non-uniform low-light images. This network leverages the feature representation of uniform natural images, employing the dark channel of the absorption light scattering model (ALSM) to estimate and enhance facial features. Furthermore, the reconstruction process is optimized through the posterior constraint of uniform image characteristics, leading to superior detail enhancement. Our approach has been rigorously tested on both synthetic and real-world images, demonstrating its effectiveness in addressing the complexities of non-uniform low-light image enhancement. The results illustrate a notable enhancement, boosting PSNR by 3.37dB and SSIM by 0.0579 compared to recent methods, with a particular focus on enhancing facial feature details. The code link is https://github.com/zbysygdsghh/face-relight.

INDEX TERMS Face image enhancement, non-uniform low-light images, deep learning, ALSM.

I. INTRODUCTION

During the process of shooting, low lighting is often encountered due to environmental factors. Particularly in complex lighting conditions, the images can exhibit non-uniform low-exposure states. For example, when facing the backlight source [1], dealing with complex light sources, or other similar situations, the critical foreground faces cannot be effectively identified, as depicted in Fig. 1(a).

Existing illumination enhancement algorithms, such as those based on histogram correction or fusion techniques, find it challenging to achieve satisfactory enhancement in such scenarios. Consequently, several enhancement methods are difficult to achieve desirable enhancement for handling these complex scenes as some approaches [2].

In addressing this issue, non-uniform illumination enhancement algorithm [3] offers the analysis of the process of the changing illumination. These algorithms incorporate adaptive implementations of multi-illumination state enhancement [4] or image lighting analysis [5] during the image enhancement process. The adaptive illumination enhancement algorithms primarily employ local and global techniques, utilizing multiple exposure intensities, to dynamically and optimally enhance the scene. This adaptive approach ensures that the image is processed in a way that best enhances the overall visual quality.

The illumination analysis approaches primarily concentrate on utilizing image filters, such as the Naturalness Preserved Enhancement algorithm (NPE) [6] and the Probabilistic Method for Image Enhancement (SEIR) [7], or deep networks to analyze pixel intensities within the image based on the tools dealing with images [8]. However, when dealing with backlit images, a notable lighting contrast between the foreground object and the background is observed. Unfortunately, in many instances, the foreground object appears severely underexposed. As shown in Fig. 1(b), there

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

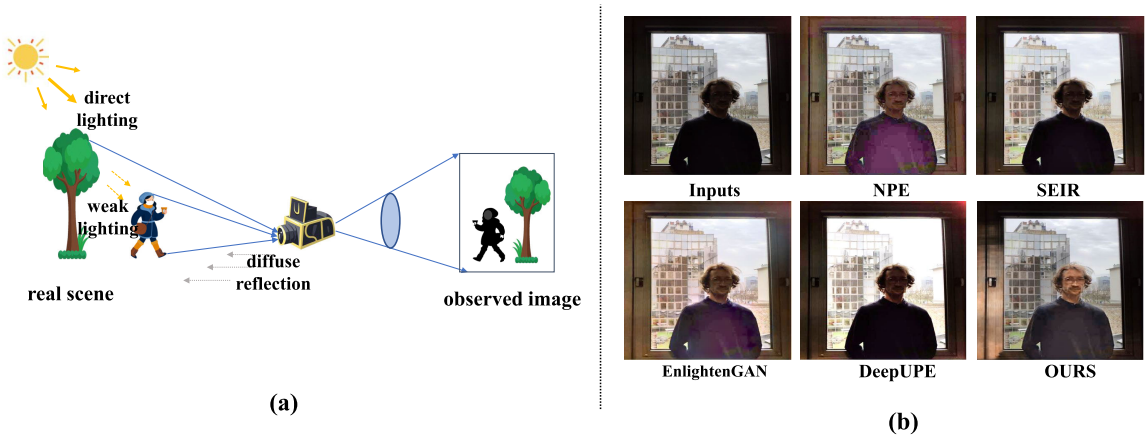

FIGURE 1. (a) The scene of backlit images. In the presence of direct sunlight, the background of the scene becomes overexposed while the foreground object appears underexposed due to glaring or diffuse reflections. Consequently, these complex lighting conditions pose challenges in capturing important information in the foreground of backlit images. (b) The results of both the compared methods and our proposed method for a real backlit image are presented. Among these methods, our approach demonstrates superior performance in enhancing the details of the foreground faces.

are several challenges that hinder effective foreground object recovery using adaptive enhancement techniques for backlit images. Firstly, the significant contrast in brightness between the foreground and background makes it difficult to achieve satisfactory foreground recovery. In some cases, the recovery process may lead to excessive exposure of the background, resulting in an unnatural appearance of the image. Secondly, the extremely low exposure of the foreground makes it challenging to restore fine details in that area. Lastly, when working with real images that lack sample labels and significant variations in lighting conditions, it becomes difficult to accurately and robustly represent the lighting state through network training. Consequently, the existing approaches prove ineffective in enhancing such backlit images.

To tackle the aforementioned challenges, we propose a non-uniform illumination enhancement algorithm that utilizes dark channel analysis under complex illumination conditions. Our approach aims to address three main problems. Firstly, by understanding the formation mechanism of backlit images, we leverage the dark channel to effectively estimate the illumination distribution of low-quality images. This allows us to achieve a more accurate representation of the image's state, even in the absence of explicit illumination labels. Secondly, we unify the model of the dark channel with a deep network to derive an enhancement algorithm. This algorithm leverages facial feature representations to reconstruct low-quality images, leading to improved overall image quality. Lastly, we combine the presentation state of the image under natural uniform illumination with the reconstructed image's posterior. We extract key features of the image using a shared network and use these features to constrain the illumination state and details of the reconstructed image, further enhancing its visual quality.

Experimental results on both synthetic and real data have shown that our method outperforms the state-of-the-art in the enhanced images.

To sum up, **contributions** of the work are as follows:

1) We analyze the model ALSM through the formation mechanism of backlit images of the actual scene, and combine the dark channel prior and the brightness analysis of the low-quality image to realize the optimization of the low-quality backlit images for face details.
2) We have effectively analyzed the dark channel rules and data statistics of backlit face images in actual scenes, summarized their presentation rules in images, and effectively applied them to the deep learning network.
3) We build a deep learning network based on the dark channel prior and the posterior constraints of the uniform face feature representation, and realize the reconstruction representation of low-quality backlit images through the loss functions.

The remainder of the paper is organized as follows. In Section II, we introduce some related works. Our method is proposed in Section III. The experiments on synthetics and real-world low-quality images are conducted in Section IV. In Section V, we conclude our work.

II. RELATED WORK

A. LOW-LIGHT ENHANCEMENT

Low-light enhancement approaches, such as those proposed by Han et al. [9] and Gao et al. [5], aim to improve the illumination and visibility of dark images. Albu et al. [10] proposed the model based on recursive filtering and contrast stretching techniques, driven by statistical measures of the image and implemented under a logarithmic image processing model for low-light enhancement. These approaches can be broadly categorized into two main types: Retinex decomposition-based methods and deep learning-based methods. Retinex-based methods, such as single-scale Retinex [11] and multi-scale Retinex [12], employ Gaussian or bilateral filters to effectively remove halo artifacts and enhance the overall image quality. Other methods focus on manipulating both

the illumination and reflectance layers to achieve more impressive enhancement results.

Deep learning methods for enhancing low-light images have been extensively studied. Wang et al. [13] reviewed the main techniques of low-light image enhancement developed over the past decades, such as frequency-domain methods, image fusion methods, defogging model methods and machine learning methods. In a recent study by Yang et al. [14], an attempt was made to explore semi-supervised learning techniques for low-light image enhancement. The work proposed a deep recursive band (DRBN) representation that served as a connection between fully supervised and unsupervised learning frameworks, thereby leveraging the advantages of both approaches. In 2023, Wang et al. [15] introduced Low-Light Transformerbased Network (LLFormer) with the axis-based multi-head self-attention and cross-layer attention fusion block for low-light enhancement. Taking into account the semantic information of different regions, Wu et al. [16] proposed a semantic-aware knowledge-guided framework (SKF) that assisted a low-light enhancement model in learning priors encapsulated in a semantic segmentation model. Wang et al. [17] proposed the low-light image enhancement framework based on virtual exposure, its exposure control parameters are adaptively generated through a statistical analysis of the low-light image, and a virtual exposure enhancer constructed by a quadratic function is applied to generate several image frames from a single input image. Li et al. [18] proposed a knowledge distillation method for low light image enhancement based on the teacher-student framework, which realizes the student network learn the knowledge of image enhancement under the supervision of ground truth images and under the guidance of the teacher network simultaneously.

To improve the application of the models, some unsupervised methods are proposed for low-light enhancement. For example, instead of directly learning an image-to-image mapping as previous work, Guo et al. introduced the unsupervised model Zero-Reference Deep Curve Estimation (Zero-DCE) [3] for image illumination enhancement. Jiang et al. [19] proposed an unsupervised generative adversarial network, dubbed EnlightenGAN, that could be trained without low/normal-light image pairs, which generalized very well on various real-world low-quality testing images. In 2023, Fu et al. [20] proposed the unsupervised Paired Low-light Image Enhancer (PairLIE) that learned adaptive priors from low-light image pairs as the Retinex theory, and applied the priors to remove inappropriate features in the raw image with a self-supervised mechanism for well enhancement.

However, real-world low-quality images are frequently subject to various complex degradation conditions. Moreover, these images often exhibit non-uniform brightness, which further complicates the task of restoring low-quality face images. As a result, effectively restoring real-world low-quality face images under these challenging degradation conditions remains a significant hurdle to overcome.

B. FACE RELIGHTING

To solve the backlit, many works [21], [22], [23], [24], [25], [26], [27] have been proposed to improve or generate the exposure state of the image. In [28], Xie et al. observed the surface normal shape model to estimate an illumination image for the relighting.

In Portrait Shadow Manipulation (PSM), Zhang et al. [29] present a computational approach that gives casual photographers some of this control, thereby poorly-lit portraits to be relit post-capture in a realistic and easily controllable way. In [30], Wang et al. formulated the single image relighting task and proposed a novel Deep Relighting Network (DeepRelight) with scene reconversion, shadow prior estimation, and re-renderered to form the required estimation under the target light source. In recent work, Hou et al. [31] proposed the differentiable algorithm for synthesizing hard shadows of the face images based on ray tracing, which was incorporated into the face relighting model for complex illumination conditions.

However, the experimental results of the existing work show that it is difficult to reconstruct ideal enhancement results by relighting algorithms due to the special nature of the backlit illumination state, or when the foreground image is in an extremely low lighting state, where the image enhancement is poor in the absence of prior knowledge.

C. NON-UNIFORM LOW-LIGHT ENHANCEMENT

Low-illumination image enhancement approaches [3], [5], [32], [33], [34], [35] amplify illumination and improve visibility of dark images. They are classified mainly into two categories: Retinex decomposition-based [12], [36], and deep learning-based [33], [37].

The above illumination enhancement methods can not be applied to the non-uniform illumination scene directly, and then some algorithms [5], [32], [38] are proposed to fix such problem. instead of directly learning an image-to-image mapping as previous work, the methods introduced the multiple exposure illumination [4] for non-uniform illumination images enhancement adaptively.

Wang et al. [32] introduced intermediate illumination in our network to associate the input with the expected enhancement result (DeepUPE), which augmented the network's capability to learn complex photographic adjustment from expert-retouched input/output image pairs. In [1], focusing on the fact that the intensity histogram of a backlit image showed a characteristic bimodal distribution, Ueda et al. proposed an image enhancement method for single backlit images using histogram specification. In [39], Li et al. proposed a low-light image enhancement method based on the degradation model to overcome the complex situations. In 2023, Wang et al. [15] introduced Low-Light Transformerbased Network (LLFormer) with the axis-based

multi-head self-attention and cross-layer attention fusion block for low-light enhancement. To seek the results with satisfied lighting, cleanliness, and realism from degraded inputs, Guo and Hu [40] present a framework named Bread for low-light enhancement, alleviating the degradation entanglement.

Recently, several learning-based approaches have shown significant improvements in general-purpose shadow detection and manipulation, as well as non-uniform low-light enhancement. For instance, the dual hierarchically aggregation network (DHAN) [41], attentive recurrent generative adversarial network (ARGAN) [42], and stacked conditional generative adversarial network (ST-CGAN) [43] were notable examples. These approaches had demonstrated enhanced performance in shadow removal tasks.

When applying high and low brightness adaptive enhancement or shadow removal techniques for image enhancement, experimental results often indicate that the adaptive optimization is inadequate. This can lead to unsatisfactory foreground enhancement or overexposure of the background. Furthermore, due to insufficient reconstruction of foreground details, effective image restoration cannot be achieved.

D. DARK CHANNEL PRIOR AND IMAGE UNIFORM POSTERIOR

As shown in Fig. 1, the backlit scene with a direct light source generates a background state of high brightness, and in some cases even in an over-exposed situation. The foreground object can only receive weak light sources and diffuse reflections due to problems, so the foreground object is in a very low light state. The light contrast between the foreground and the background also makes the state of the foreground pixels, and it is almost impossible to identify the detailed information.

The atmospheric scattering model (ASM) focuses on the dehaze [44], [45], [46] and enhancement [47], [48], that is

$$I(x) = A \cdot R(x)t(x) + A(1 - t(x)), \tag{1}$$

where I is the observed low-quality image, and R is the reflection of the scene, with x representing the pixels in the image. A is the atmospheric light intensity, which can be seen as the constant value. t is the light transmission parameter, used to express the illumination state in our work. Inspired by ASM, Wang et al. [49] proposed the model, ALSM for low-light enhancement. In ALSM, $J(x) = A \cdot R(x)$ describes the portion of atmospheric light that is directly reflected from the scene, and it also can be seen as the target image. In a backlit environment, t forms a special distribution because the foreground object is affected by factors such as low light and diffuse reflection. In Fig 1(a), the face is in the low exposure due to the weak lighting source and lighting absorption. Then we apply the model to the uniform low-light face enhancement.

Inspired by the dark channel prior [45], in most outside scenes, the dark channel tends to zero:

$$J^{dark}(x) = \min_{y \in \Omega(x)} \left(\min_{c \in \{R,G,B\}} J^c(y) \right) \to 0, \tag{2}$$

where J^c is each channel of the color image, and $\Omega(x)$ denotes a window centered on pixel x, representing the small area around pixel point p in the window.

However, as argued, its statistical dark channel rules do not apply to the blue sky, sea, or certain situations with high exposure to direct light. The low-exposure foreground objects still meet the statistical rules of the dark channel. We have performed statistical analysis on a large amount of data and found that there is a significant difference between the dark channel statistics of the foreground and the background in Fig. 4. The dark channel provides maximum classification of the portion with different absorption rates.

Given this observation, we construct the data state of the image from the dark channel, following

$$I^{dark}(x) = J^{dark}(x)t(x) + A(1 - t(x)), \tag{3}$$

where I^{dark} and J^{dark} are respectively the dark channel priors of the low-quality and target images. The generation of t is constrained from the dark channel state such that the foreground state of the input dark channel I^{dark} converges to the highlighted state of its background, as formulated by

$$t(x) = (I^{dark}(x)/A - 1)/(J^{dark}(x)/A - 1), \tag{4}$$

and then t is applied to the atmospheric scattering model as Eq.(1) to generate the optimal image result $\hat{J}$.

To maintain the details of the reconstruction results, we introduce face features as the constraint. Since both the low-quality image I and the target image J are information interpretations of the same object, we try to introduce their common feature, the face parsing map M as an adequate information constraint for the enhancement. The facial details are extracted from the low-quality image as the distribution function:

$$p(J_r(x), M|I(x)) = p(J_r(x)|M, I(x))p(M|I(x)), \tag{5}$$

where J_r can be seen as the more information of the face details for the enhanced result $\hat{J}$. With the help of the network, we can predict precise target details J_r from the correlated information $p(J_r(x), M|I(x))$.

In addition, after effectively achieving low-quality image enhancement, we want to maintain the natural uniform state of the enhanced image $\hat{J}$ as well as to fully express the foreground detail information. As the high-quality face image in the natural state can estimate the clear face parsing map [50] for the face details, the generation $\hat{J}$ also can provide the detailed information for face features M. The distribution function is shown as

$$p(M) = p(M|H(x))p(H(x)). \tag{6}$$

where H is the high-quality image. $p(M|H(x))$ represents the common process to estimate features from the image,

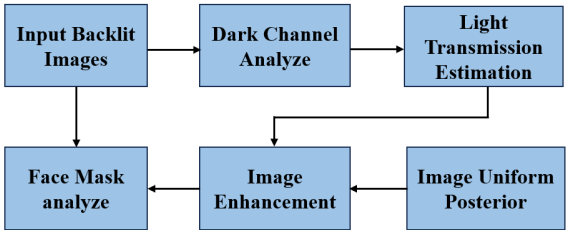

FIGURE 2. The block diagram of the our proposed framework.

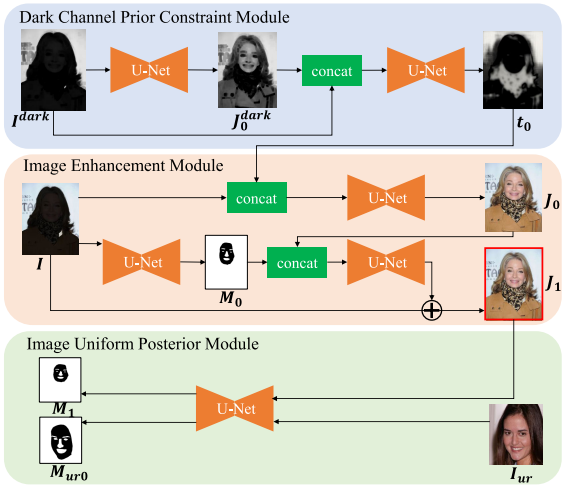

FIGURE 3. The framework of our network consists of dark channel prior constraint, image enhancement, and image uniform posterior. Firstly, we analyze the dark image with the dark channel to output the light transmission parameter t_0. Then, with t_0, we constrain the generation J_0 as the intermediate result. Following that, we feed J_0 and the estimated facial mask M_0 into the network for the result J_1. And it is the final result in the red box. Finally, we constrain the final result with the image uniform posterior. The basic network module is U-Net, and the numbers before and after U-Net indicate the number of channels for the input and output of the module.

and then this process can be applied to the enhanced image $\hat{J}$. If the facial features can be estimated from the enhanced image by this common process, it proves that the generation $\hat{J}$ has adequate and effective details to approach the target one in a uniform state. Rather than discriminative networks, [51] that introduce unexpected artifacts, we prefer to improve the accuracy and uniform brightness by using the feature representation as the posterior constraint.

In Fig. 2, we present the block diagram of the our proposed framework.

III. PROPOSED METHOD

A. NETWORK

Our solution is mainly based on the dark channel rules in the backlit scenario, and the main algorithm is implemented through end-to-end network architecture. In this section, we present its framework. According to the analyzed statistical rules, the network structure for image enhancement mainly consists of the following parts: one is the dark channel prior constrained network for illumination analysis and the generation of enhancement, and another part is for image enhancement under illumination analysis and rule constraints. Meanwhile, the posterior constraint network is provided to better optimize the uniform representation of the foreground and background of low-quality images.

1) DARK CHANNEL PRIOR CONSTRAINT MODULE

The dark channel prior constraint module is mainly used for image brightness analysis and dark channel generation constraints. The module mainly consists of two parts. A process is used to generate the target dark channel J^{dark} from the input dark channel I^{dark}, as shown in

$$J_0^{dark}(x) = U_0(I^{dark}(x)), \tag{7}$$

where U_0 denotes the network, and J_0^{dark} denotes the dark channel output by the training.

Another process is to combine the input dark channel I^{dark} and the generated output J_0^{dark} to predict the illumination parameters under the dark channel constraint by a network, formally

$$t_0(x) = U_1(concat(I^{dark}(x), J_0^{dark}(x))), \tag{8}$$

where U_1 denotes the generative network and *concat* is used to concatenate these feature maps together. t_0 is the output of the light parameters predicted by the network. The dark channel constraint is mainly used to predict the illumination parameters to constrain image enhancement.

2) IMAGE ENHANCEMENT MODULE

The image enhancement module mainly targets the enhancement of the image, which is based on the dark channel prior constraint. To better reconstruct the low-quality foreground objects of the image, we introduce face features as detailed optimization and further enhancement constraints in the enhancement process. The module consists of three parts. The first process focuses on the input of the low-quality image I, generating the face feature map M:

$$M_0(x) = U_2(I(x)), \tag{9}$$

where U_2 denotes the generation network, and M_0 denotes the face feature output. The second process is to enhance the input low-quality image based on the illumination parameters of the dark channel to generate the face enhancement result J:

$$J_0(x) = U_3(concat(I(x), t_0(x))), \tag{10}$$

where U_3 denotes the generation network, and J_0 denotes the enhanced image output. The third process is based on the foreground object feature map constraint to further enhance the image and generate the result of further enhancement of the face:

$$J_1(x) = I(x) + U_4(concat(M_0(x), J_0(x))), \tag{11}$$

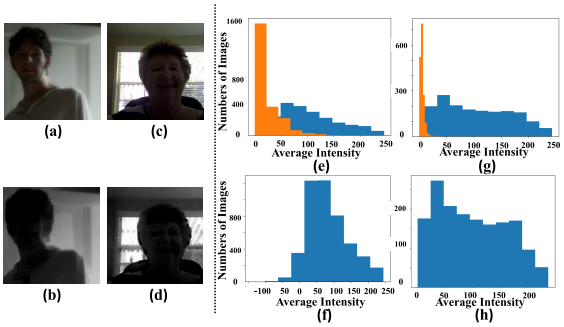

FIGURE 4. The listed real backlit images (a),(c) and their corresponding dark channel images (b),(d). (e) is the statistical average intensity of foreground and background for the about 2000 selected images in actual scenes, where blue denotes the background and orange indicates the foreground. (f) shows the number of different average values between the background and foreground of the same image sequence. (g) is the statistical average intensity of foreground and background for about 2000 synthetic images. (h) shows the number of different average values between the background and foreground of the same sequence.

where U_4 denotes the generated network, J_1 denotes the enhanced image output, expressed jointly by I and residual features. J_1 is also used as the final expression of the network enhancement.

Further, to better make the enhancement result J_1 in a uniform illumination state, we introduce a uniform expression posterior module to constrain our reconstruction results by a large number of uniform illumination results.

3) IMAGE UNIFORM POSTERIOR MODULE

The image uniform posterior module is mainly used to constrain the results of our reconstruction enhancement by the generated results of other uniformly illuminated images. The network module includes a generative network, which is a parameter-sharing network structure. The input-output pairs of shared parameters include our enhanced result J_1 as input and its generated object feature map M_1 as output. The object training set I_{ur}, which is independent of the sample set, is used for input and its generated feature map M_{ur} as output, as follows:

$$M_1(x) = U_5(J_1(x)), M_{ur0}(x) = U_5(I_{ur}(x)), \tag{12}$$

where U_5 denotes the generative network and it provide the sharing weighting for the images $J_1(x)$ and $I_{ur}(x)$. M_{ur0} denotes the foreground object feature maps from input I_{ur}.

B. LOSS FUNCTIONS

The setting of the loss function also mainly refers to the design of the network structure, which mainly includes three parts: the prior loss function for the dark channel, the loss function for image enhancement, and the posterior loss function for uniform expression.

The dark channel prior loss function mainly constrains the generation of two outputs, the target dark channel J_0^{dark} and the luminance parameter t_0 as

$$\begin{aligned} Loss_0 = &||J_0^{dark}(x) - J^{dark}(x)||_2^2 \\ &+ ||(I^{dark}(x)/A - 1) - (J^{dark}(x)/A - 1)t_0(x)||_2^2 \\ &+ ||(I(x)/A - 1) - (J(x)/A - 1)t_0(x)||_2^2. \end{aligned} \tag{13}$$

where A can be estimated as a constant value by [45] for the loss function. And t_0 is the feature parameters to constrain Eq. 1 and Eq. 3 meanwhile.

The image enhancement loss function is used for the enhancement of the final image results, constraining two outputs, such as object feature maps M_0 and the image enhancement output J_1.

$$Loss_1 = ||M_0(x) - M(x)||_2^2 + ||J_1(x) - J(x)||_2^2. \tag{14}$$

The uniform posterior loss function is used to constrain the uniform expression of the image, including the object feature map M_1 of the enhancement result and the feature map M_{ur0} of the posterior dataset, whose functions are as follows:

$$Loss_2 = ||M_1(x) - M(x)||_2^2 + ||M_{ur0}(x) - M_{ur}(x)||_2^2. \tag{15}$$

The overall loss function is the combination of the three loss functions:

$$Loss_t = \alpha Loss_0 + \beta Loss_1 + \gamma Loss_2. \tag{16}$$

where α,β,γ are used to balance the overall loss function.

C. IMPLEMENTATION DETAILS

All the network frameworks used for feature extraction and image generation are based on U-Net [37], [52], mainly because the network structure has fewer parameters compared to residual networks (ResNet) [53], dense networks (DenseNet) [54], etc. The intermediate skip connection allows for more efficient loss gradient transfer, thus providing a more comprehensive reconstruction performance compared to other network frameworks. The model is trained by ADAM optimizer [55]. The balance parameters in the overall loss function are respectively $\alpha = 0.001, \beta = 1, \gamma = 0.01$. The learning parameter is set to 0.0001, where the parameter value is halved every 100 epochs during training. The total number of epochs is 400, with a final learning rate set to $1.25e{-}5$. Our experiments are implemented by Pytorch [56] on an NVIDIA RTX 1080ti GPU.

IV. EXPERIMENTS

We experimentally verified the effectiveness of our algorithm on backlit images. As a special scene of low-quality images, the contrast algorithms include two types. One can be seen as non-uniform low light enhancement and the other can be seen as shadow removal. The contrast non-uniform low light enhancement algorithms include the classical traditional algorithms NPE [6], SEIR [7], the SOTA framework of deep learning algorithms, DeepUPE [32], PairLIE [20], LLFormer [15] and the unsupervised methods Zero-DCE [3] and EnlightenGAN [19]. Shadow removal algorithms include the recent algorithms DeepRelight [30]

TABLE 1. The objective metrics (PSNR and SSIM [60]) of the ablation experiments w/o DC, w/o UF, w/o DCUF, and ours.

	w/o DC	w/o UF	w/o DCUF	OURS
PSNR(db)	24.49	24.69	19.33	**28.16**
SSIM	0.8322	0.8186	0.7552	**0.8910**

and ST-CGAN [43]. All compared deep learning algorithms use provided pre-trained models for evaluation on various simulated and real datasets.

A. DATESETS

The simulation data used for training is from the dataset CelebA, where we extract the foreground M_s, and the background information M_b by the matting algorithm:

$$I = M_b + \theta M_s, \tag{17}$$

where θ is the balance parameter of the foreground M_s and background M_b, and its value range is set to a random number between $[0.1, 0.2]$. The input image is simulated and the original image is used as the target image for training. All training data are tuned to 512×512. In addition, the training data used to uniformly express the posterior constraints during training are CelebAMask-HQ [57] and its mask data that does not overlap with the training test data. The simulation data used for testing is processed by similar simulations, such as CelebA [58] and Helen Dataset [59]. In addition, about 2500 backlit images are extracted from the web as test data for the actual scene.

To count the dark channel rule of the backlit data, we do statistical analysis of the data on more than 2000 images randomly selected from actual and synthetic scenes as Fig. 4. The images in Fig. 4(a) and (c) are the low-quality inputs, and the images in Fig. 4(b) and (d) represent the corresponding dark channel maps as [45]. As we have discussed, the background of the backlit image presents large values of pixels because it is directly exposed to the light source, so the background is in a high pixel state compared to the dark channel rule of the foreground object. In contrast, in Fig. 4(e), we enumerate the comparison of the average values of the foreground and background for the statistics, with blue as the background and orange as the foreground pixels. The two have contrasting ranges of values and are in significantly different distributions. In Fig. 4(f), we also show the distribution of the difference in the average values of the foreground and background of the same image, with values mainly distributed between $[50, 150]$, indicating that the foreground and background of the same image can mostly be clearly distinguished by the dark channel prior. The presented images and statistics demonstrate the validity of the dark channel priors for the analysis of the backlit images, or the low-quality state of these images. Fig. 4(g) and (h) also show the corresponding distribution of selected 2000 synthetic images, which are similar to the distribution of the image in Fig. 4(e) and (f) in reality.

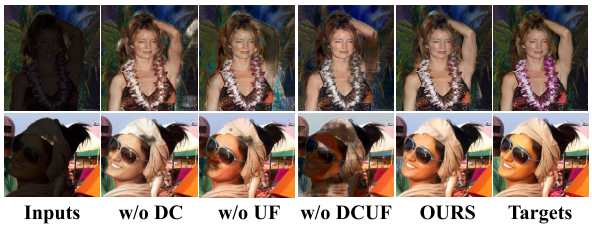

FIGURE 5. The generated images of the ablation experiments w/o DC, w/o UF, w/o DCUF, and ours.

B. ABLATION STUDY

In order to verify the respective contributions of the dark channel prior, image enhancement, and uniform expression posterior constraints to the image reconstruction, we perform ablation experiments of the respective modules. These include making $\alpha = 0$ to remove the dark channel constraint (w/o dark channel as w/o DC), making $\gamma = 0$ to remove the uniform constraint (w/o uniform as w/o UF), and making $\alpha = 0$ and $\gamma = 0$ to remove the prior and posterior constraints (w/o dark channel and uniform as w/o DCUF).

In Fig. 5, the experimental results without dark channel and uniform constraint (w/o DCUF) are the worst ones, and the reconstructed images still suffer from severe non-uniform low illumination in the absence of dark channel and uniform image constraints. In the case of missing uniform image constraint (w/o UF), the reconstructed foreground image shows poor color state and poor reconstruction details, which proves the effect of uniform image posterior constraint on the lighting and color of the reconstructed image. And the reconstruction results in the absence of dark channel analysis (w/o DC) show that the non-uniform low-light images still affect the details of the foreground objects, proving the contribution of the dark channel in the network training process.

The objective metrics in Tab. 1 also demonstrate and validate the effectiveness of the respective modules for the final experimental results. Among them, the objective metrics are the smallest w/o DCUF, which can be viewed as the worst reconstruction performance in the absence of constraints. w/o DC and w/o UF, the metrics of PSNR and SSIM have a slight deviation, but both are small compared to the final reconstruction results, proving the effectiveness of our designed framework.

Analysis of Parameters: We also analyze the effects of the different parameters α, β and γ. Since the parameters have the balancing ratio, we always set $\beta = 1$ to test the impact of different values of α and γ on the experimental metrics PSNR. As shown in the Fig. 6, the optimal parameters are $\alpha = 0.01$ and $\gamma = 0.001$. With the increase of parameters, their performance gradually decreases with a slight trend.

State Analysis of Constraint Generations: In Fig. 7, we list several images of intermediate results. M_0 and M_1 represent the features of the foreground objects generated by the input image reconstruction and uniform constraint, respectively, where M_1 is better rendered and the results show the performance of the generated enhanced results. The input and

TABLE 2. Average PSNR(db), SSIM, NIQE [61] and il-NIQE [62] of our results and the comparison methods NPE [6], SEIR [7], DeepUPE [32], EnlightenGAN [19], Zero-DCE [3], LLFormer [15]+DC, Bread [40]+DC, and ST-CGAN [43]. Par denotes the parameters of the methods, Flops denotes the floating point operations, and Time denotes the running time of the methods.

Algorithms	CelebA				Helen				Par	Flops	Time
	↑ PSNR	↑ SSIM	↓ NIQE	↓ il-NIQE	↑PSNR	↑SSIM	↓NIQE	↓ il-NIQE			
NPE	14.88	0.7300	15.92	45.39	12.72	0.6872	15.93	45.16	/	/	3.76s
SEIR	12.49	0.6455	15.30	49.02	14.05	0.6536	15.55	47.54	/	/	0.247s
DeepUPE	16.08	0.7122	15.39	50.95	16.54	0.6561	15.36	47.16	0.43M	0.13G	0.021s
EnlightenGAN	14.83	0.7273	15.50	44.68	15.21	0.5604	15.67	**44.89**	8.64M	11.82G	0.010s
Zero-DCE	15.71	0.7564	15.72	44.98	16.09	0.7042	16.03	45.98	0.079K	3.97G	0.003s
LLFormer+DC	24.79	0.8331	15.48	54.31	23.38	0.6976	**14.86**	54.31	24.5M	34.60G	0.082s
Bread+DC	23.42	0.8308	15.41	59.41	22.84	0.6921	14.91	51.02	8.21M	50.74G	0.036s
ST-CGAN	13.26	0.6213	18.24	50.21	12.73	0.4491	18.24	50.28	58.5M	36.05G	0.169s
Ours	**28.16**	**0.8910**	**15.20**	**44.17**	**27.12**	**0.7965**	15.13	45.12	11.7M	96.91G	0.034s

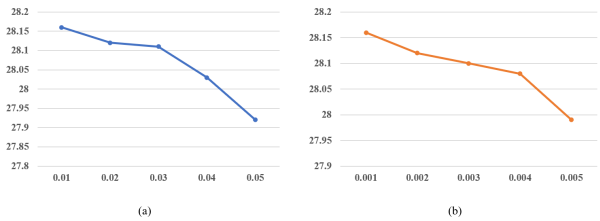

FIGURE 6. The ablation study of the different parameters.(a) The metrics PSNR of the different values of the parameter α, (b) the metrics PSNR of the different values of the parameter γ.

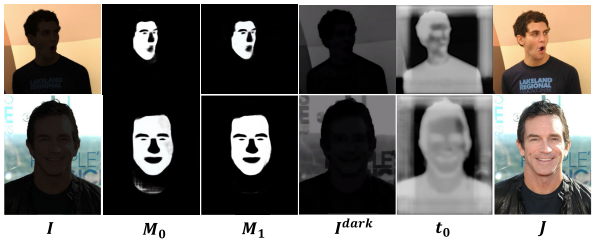

FIGURE 7. The intermediate generations during the training. The images are corresponding to the feature maps in the modules of image reconstruction and uniform posterior. The images of M_0 and M_1 are the object features in Eq. (14) and Eq. (15). The images of I, I^{dark}, t_0 and J are related to Eq. (13).

output maps are shown as I and J in the figure. while I^{dark} is the image result of the dark channel. The maps t_0 in the figure as Eq. (8) are mainly used as an important parameter of the dark channel to constrain the generation of final results. The presented results also reflect difference in brightness between face and background.

C. COMPARISONS WITH COMPETITIVE METHODS

In this section, we compare our results with the compared methods. Since our work aims at recovering low-quality backlit images, the performance is verified on simulated datasets CelebA and Helen. The comparison experiments are tested mainly on our training model as well as on the existing comparison algorithms. These include the non-uniform luminance enhancement algorithms NPE, SEIR, DeepUPE, Zero-DCE, and EnlightenGAN, as well as the de-shadowing algorithms DeepRelight and ST-CGAN. To address the issue of a dark foreground in the results obtained using pre-trained LLFormer and Bread, we incorporated the constraints of our proposed dark channel prior during training. They are denoted as LLFormer [15]+DC and Bread [40]+DC.

As shown in Fig. 8, the input images are simulated from the Helen (first row) and CelebA (second through fourth rows). Similar to the real backlit image, the foreground object in the figure is in extreme darkness and almost unrecognizable, while the background is in normal light or even over-exposed state. The reconstruction results of the traditional non-uniform illumination enhancement algorithm NPE can achieve some enhancement effects in the foreground, but still in low light, while the background image is somewhat overexposed, and the enhancement effect of SEIR for foreground objects is poor compared with NPE, as some images are still in low light. Deep learning non-uniform illumination enhancement algorithms DeepUPE, Zero-DCE, and EnlightenGAN all have some enhancement effects on the input image. Zero-DCE has the best effect, but the foreground object is still shimmering and the background color is distorted; EnlightenGAN has better foreground enhancement, but the background is overexposed; DeepUPE has the worst enhancement effect, where the main problem is the poor brightness enhancement of the foreground and the most serious overexposure of the background. The de-shadowing algorithms DeepRelight and ST-CGAN both have serious problems with the enhancement of the input and poor image recovery. Compared to all comparison algorithms, our method can effectively achieve foreground recovery while maintaining as much background image information as possible, which is mainly dependent on image analysis in the dark channel and a posteriori constraints on uniform image reconstruction. In contrast, in the absence of effective image resolution, NPE, DeepUPE, and EnlightenGAN have some overexposure in the background during image enhancement. And SEIR and Zero-DCE are ineffective for foreground enhancement. And shadow removal algorithms DeepRelight and ST-CGAN are poor for face enhancement, largely because there is some difference between the imaging method of backlit images and shadow occlusion, or the face of backlit images is in low light. The introduction of the dark channel prior for the methods LLFormer and Bread can realize the good visual performance of the backlit images.

FIGURE 8. The simulated results of compared methods and ours. The comparison methods include NPE, SEIR, DeepUPE, Zero-DCE, EnlightenGAN, DeepRelight, and ST-CGAN. The images in the first row are from the Helen dataset, and the results in the following rows are from the CelebA dataset.

FIGURE 9. (a) The results of real backlit images generated by comparison methods and ours.(b) The limitations for our results of real backlit images, (b1) and (b2) represent different facial images.

The de-shadowing algorithm ST-CGAN has serious problems with the enhancement of the input and poor image recovery. Compared to all comparison algorithms, our method can effectively achieve foreground recovery while maintaining as much background image information as possible, which is mainly dependent on image analysis in the dark channel and a posteriori constraints on uniform image reconstruction.

As in Tab. 2, we list several main objective metrics, which include PSNR, SSIM [60], and the no-reference performance metrics NIQE [61], il-NIQE [62]. According to the metrics, when compared with other methods, our approach improves PSNR by 3.37dB and SSIM by 0.0579, while decreasing NIQE by 0.1 and il-NIQE by 0.51 on the synthetic CelebA dataset. On the synthetic Helen dataset, our method enhances PSNR by 3.74dB and SSIM by 0.0923 compared to the methods. However, our method ranks third in terms of NIQE and second in terms of il-NIQE metrics. For optimal outcomes in most scenarios, this affirms the efficacy of our method and the robustness of the model, validated both subjectively and objectively. We also list the compared computation complexity, and ours are at a moderate level in terms of parameters, Flops (floating point operations) and running time.

D. REAL-WORLD IMAGES

We have also evaluated our method on real-world images. Our results are recovered from the models trained from synthetic low-light images of CelebA datasets. As shown in Fig. 9, frontal and non-frontal facial backlit images are captured from cameras. The results of SEIR and DeepUPE suffer from the low-light conditions. The results of Zero-DCE and EnlightenGAN have better enhancement performance, however, their backgrounds are over-exposed and the foreground objects are unclear enough. Compared to the above methods, our results enjoy the most detailed information and natural illumination.

E. LIMITATIONS FOR REAL-WORLD IMAGES

In Fig. 9(b), our method has some limitations for the real-world images. As in Fig. 9(b1), the captured image is affected

by extremely dark foreground objects, then our method does not perform well on such low-quality images for more details. Then the result is also hard to recognize. As in Fig. 9(b2), foreground object, such as faces, is also affected by uniform illumination. Although our method can achieve brightness enhancement, it does not completely maintain the uniform brightness, even for foreground objects, mainly because our constraints and simulation data do not take into account the uniform illumination of foreground objects, so it is impossible to keep image brightness uniform for testing, which is the direction for optimization.

V. CONCLUSION

Our research addresses the challenge of low-quality backlit images, particularly those with low-light foreground objects. We investigate the underlying principles of the dark channel within ALSM for backlit images. Building upon this analysis, we propose a deep reconstruction network that incorporates these principles and integrates feature representations from uniform natural images to impose posterior constraints. Extensive experimentation demonstrates the superiority of our approach over recent methods, resulting in a notable increase in PSNR by 3.37dB and SSIM by 0.0579. However, our method encounters limitations when applied to real-world images with extreme darkness. Additionally, due to the scarcity of paired data for non-uniform low-light face images, we aim to develop an unsupervised method utilizing brightness analysis models for image enhancement in future research.

CONFLICT OF INTEREST STATEMENTS

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of the manuscript.

REFERENCES

[1] Y. Ueda, D. Moriyama, T. Koga, and N. Suetake, "Histogram specification-based image enhancement for backlit image," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2020, pp. 958–962.

[2] Y. Wang, R. Wan, W. Yang, H. Li, L.-P. Chau, and A. C. Kot, "Low-light image enhancement with normalizing flow," in *Proc. AAAI Conf. Artif. Intell.*, 2022, pp. 2604–2612.

[3] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, and R. Cong, "Zero-reference deep curve estimation for low-light image enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1777–1786.

[4] F. Lv, B. Liu, and F. Lu, "Fast enhancement for non-uniform illumination images using light-weight CNNs," in *Proc. 28th ACM Int. Conf. Multimedia*, Oct. 2020, pp. 1450–1458.

[5] Y. Gao, H.-M. Hu, B. Li, and Q. Guo, "Naturalness preserved nonuniform illumination estimation for image enhancement based on retinex," *IEEE Trans. Multimedia*, vol. 20, no. 2, pp. 335–344, Feb. 2018.

[6] 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.

[7] X. Fu, Y. Liao, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding, "A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 4965–4977, Dec. 2015.

[8] L. Alzubaidi, J. Bai, A. Al-Sabaawi, J. Santamaría, A. S. Albahri, B. S. N. Al-Dabbagh, M. A. Fadhel, M. Manoufali, J. Zhang, A. H. Al-Timemy, Y. Duan, A. Abdullah, L. Farhan, Y. Lu, A. Gupta, F. Albu, A. Abbosh, and Y. Gu, "A survey on deep learning tools dealing with data scarcity: Definitions, challenges, solutions, tips, and applications," *J. Big Data*, vol. 10, no. 1, p. 46, Apr. 2023.

[9] X. Han, H. Yang, G. Xing, and Y. Liu, "Asymmetric joint GANs for normalizing face illumination from a single image," *IEEE Trans. Multimedia*, vol. 22, no. 6, pp. 1619–1633, Jun. 2020.

[10] F. Albu, C. Vertan, C. Florea, and A. Drimbarean, "One scan shadow compensation and visual enhancement of color images," in *Proc. 16th IEEE Int. Conf. Image Process. (ICIP)*, Nov. 2009, pp. 3133–3136.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] T. Wang, K. Zhang, T. Shen, W. Luo, B. Stenger, and T. Lu, "Ultra-high-definition low-light image enhancement: A benchmark and transformer-based method," in *Proc. AAAI Conf. Artif. Intell.*, vol. 37, no. 3, 2023, pp. 2654–2662.

[16] Y. Wu, C. Pan, G. Wang, Y. Yang, J. Wei, C. Li, and H. T. Shen, "Learning semantic-aware knowledge guidance for low-light image enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 1662–1671.

[17] W. Wang, D. Yan, X. Wu, W. He, Z. Chen, X. Yuan, and L. Li, "Low-light image enhancement based on virtual exposure," *Signal Process., Image Commun.*, vol. 118, Oct. 2023, Art. no. 117016.

[18] Z. Li, Y. Wang, and J. Zhang, "Low-light image enhancement with knowledge distillation," *Neurocomputing*, vol. 518, pp. 332–343, Jan. 2023.

[19] 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.

[20] Z. Fu, Y. Yang, X. Tu, Y. Huang, X. Ding, and K.-K. Ma, "Learning a simple low-light image enhancer from paired low-light instances," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 22252–22261.

[21] P. Choudhary, S. Mannar, A. Kumar, and M. Son, "Real-time face relighting via adaptive normal mapping," in *Proc. SIGGRAPH Asia Mobile Graph. Interact. Appl.*, Nov. 2017, pp. 1–6.

[22] Z. Shu, S. Hadap, E. Shechtman, K. Sunkavalli, S. Paris, and D. Samaras, "Portrait lighting transfer using a mass transport approach," *ACM Trans. Graph.*, vol. 36, no. 4, p. 1, Jul. 2017.

[23] O. Aldrian and W. A. P. Smith, "Inverse rendering of faces with a 3D morphable model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 5, pp. 1080–1093, May 2013.

[24] Y. Wang, L. Zhang, Z. Liu, G. Hua, Z. Wen, Z. Zhang, and D. Samaras, "Face relighting from a single image under arbitrary unknown lighting conditions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 11, pp. 1968–1984, Nov. 2009.

[25] H. Zhou, S. Hadap, K. Sunkavalli, and D. Jacobs, "Deep single-image portrait relighting," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 7193–7201.

[26] T. Nestmeyer, J.-F. Lalonde, I. Matthews, and A. Lehrmann, "Learning physics-guided face relighting under directional light," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 5123–5132.

[27] K. Kaur, N. Jindal, and K. Singh, "Fractional Fourier transform based Riesz fractional derivative approach for edge detection and its application in image enhancement," *Signal Process.*, vol. 180, Mar. 2021, Art. no. 107852.

[28] X. Xie, J. Lai, C. Y. Suen, and W.-S. Zheng, "Non-ideal class non-point light source quotient image for face relighting," *Signal Process.*, vol. 91, no. 4, pp. 1048–1053, Apr. 2011.
[29] X. Zhang, J. T. Barron, Y.-T. Tsai, R. Pandey, X. Zhang, R. Ng, and D. E. Jacobs, "Portrait shadow manipulation," *ACM Trans. Graph.*, vol. 39, no. 4, pp. 1–78, Aug. 2020.
[30] L.-W. Wang, W.-C. Siu, Z.-S. Liu, C.-T. Li, and D. P. Lun, "Deep relighting networks for image light source manipulation," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*. Cham, Switzerland: Springer, 2020, pp. 550–567.
[31] A. Hou, M. Sarkis, N. Bi, Y. Tong, and X. Liu, "Face relighting with geometrically consistent shadows," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 4207–4216.
[32] 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. 6842–6850.
[33] 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. 3060–3069.
[34] X. Ding, Z. Wang, J. Fang, Z. Shu, R. Hu, and C.-W. Lin, "Watch you under low-resolution and low-illumination: Face enhancement via bi-factor degradation decoupling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 34, no. 5, pp. 3481–3495, May 2024.
[35] S. Zhang and E. Y. Lam, "An effective decomposition-enhancement method to restore light field images captured in the dark," *Signal Process.*, vol. 189, Dec. 2021, Art. no. 108279.
[36] C. Xiao and Z. Shi, "Adaptive bilateral filtering and its application in retinex image enhancement," in *Proc. 7th Int. Conf. Image Graph.*, Jul. 2013, pp. 45–49.
[37] C. Chen, Q. Chen, J. Xu, and V. Koltun, "Learning to see in the dark," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3291–3300.
[38] A. Mustapha, A. Oulefki, M. Bengherabi, E. Boutellaa, and M. A. Algaet, "Towards nonuniform illumination face enhancement via adaptive contrast stretching," *Multimedia Tools Appl.*, vol. 76, no. 21, pp. 21961–21999, Nov. 2017.
[39] P. Li, J. Liang, and M. Zhang, "A degradation model for simultaneous brightness and sharpness enhancement of low-light image," *Signal Process.*, vol. 189, Dec. 2021, Art. no. 108298.
[40] X. Guo and Q. Hu, "Low-light image enhancement via breaking down the darkness," *Int. J. Comput. Vis.*, vol. 131, no. 1, pp. 48–66, Jan. 2023.
[41] J. Wang, X. Li, and J. Yang, "Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 1788–1797.
[42] B. Ding, C. Long, L. Zhang, and C. Xiao, "ARGAN: Attentive recurrent generative adversarial network for shadow detection and removal," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 10212–10221.
[43] X. Cun, C.-M. Pun, and C. Shi, "Towards ghost-free shadow removal via dual hierarchical aggregation network and shadow matting GAN," in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 10680–10687.
[44] S. K. Nayar and S. G. Narasimhan, "Vision in bad weather," in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, Sep. 1999, pp. 820–827.
[45] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341–2353, Dec. 2011.
[46] R. Fattal, "Single image dehazing," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 1–9, 2008.
[47] 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.
[48] Z. Gu, C. Chen, and D. Zhang, "A low-light image enhancement method based on image degradation model and pure pixel ratio prior," *Math. Problems Eng.*, vol. 2018, pp. 1–19, Jul. 2018.
[49] Y.-F. Wang, H.-M. Liu, and Z.-W. Fu, "Low-light image enhancement via the absorption light scattering model," *IEEE Trans. Image Process.*, vol. 28, no. 11, pp. 5679–5690, Nov. 2019.
[50] B. M. Smith, L. Zhang, J. Brandt, Z. Lin, and J. Yang, "Exemplar-based face parsing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 3484–3491.
[51] I. J. Goodfellow, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020.
[52] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, vol. 9351. Cham, Switzerland: Springer, 2015, pp. 234–241.
[53] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[54] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.
[55] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[56] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in Pytorch," in *Proc. Neural Inf. Process. Syst. (NeurIPS)*, 2017.
[57] C.-H. Lee, Z. Liu, L. Wu, and P. Luo, "MaskGAN: Towards diverse and interactive facial image manipulation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 5548–5557.
[58] Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep learning face attributes in the wild," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 3730–3738.
[59] V. Le, J. Brandt, Z. Lin, L. Bourdev, and T. S. Huang, "Interactive facial feature localization," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 679–692.
[60] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[61] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.
[62] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2579–2591, Aug. 2015.

BO-YU ZHANG is currently pursuing the bachelor's degree in communication engineering with the School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou, China. His research interests include object detection and segmentation and machine learning.

QING-CHUN ZHANG is currently pursuing the bachelor's degree in communication engineering with the School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou, China. Her research interests include image processing and machine learning.

WEN-YING ZHANG is currently a Lecturer with the School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou, China. Her research interests include image processing and pattern recognition.