ALL-E: Aesthetics-guided Low-light Image Enhancement

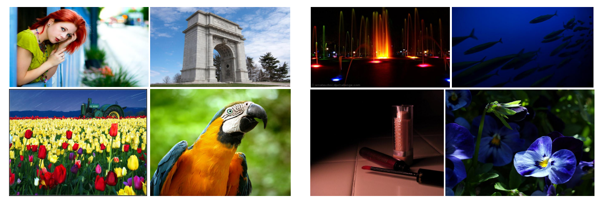
Ling Li , Dong Liang , Yuanhang Gao , Sheng-Jun Huang , Songcan Chen
College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics
MIIT Key Laboratory of Pattern Analysis and Machine Intelligence
{liling, liangdong, gaoyuanhang, huangsj, s.chen}@nuaa.edu.cn

Abstract

Evaluating the performance of low-light image enhancement (LLE) is highly subjective, thus making integrating human preferences into image enhancement a necessity. Existing methods fail to consider this and present a series of potentially valid heuristic criteria for training enhancement models. In this paper, we propose a new paradigm, *i.e.*, aesthetics-guided low-light image enhancement (ALL-E), which introduces aesthetic preferences to LLE and motivates training in a reinforcement learning framework with an aesthetic reward. Each pixel, functioning as an agent, refines itself by recursive actions, *i.e.*, its corresponding adjustment curve is estimated sequentially. Extensive experiments show that integrating aesthetic assessment improves both subjective experience and objective evaluation. Our results on various benchmarks demonstrate the superiority of ALL-E over state-of-the-art methods. Source code: https://dongl-group.github.io/project_pages/ALLE.html

1 Introduction

[1]Due to the limitations inherent in optical devices and the variability of external imaging conditions, images are often captured with poor lighting, under-saturation, and a narrow dynamic range. These degradation factors significantly impair the visual aesthetics of the images, and lead to detrimental effects on a wide range of downstream computer vision and multimedia tasks [Yu *et al.*, 2021b; Cho *et al.*, 2020; Guo *et al.*, 2020]. Manually editing and enhancing low-light images is time-consuming, even for a professional. Therefore, many learning-based strategies have been proposed, introducing a series of potentially valid heuristic constraints for training image enhancement models [Lore *et al.*, 2017; Wei *et al.*, 2018; Zhang *et al.*, 2019; Xu *et al.*, 2020; Ren *et al.*, 2019; Liang *et al.*, 2022]. However, the aforementioned methods fail to account for the importance of subjective human evaluation in low-light image enhancement (LLE) tasks. On an online professional photography contest – DPChallenge[2], the subjective human aesthetic scores of normal images are much higher than those of low-light images, and the former provides a significantly superior visual experience compared to the latter, as shown in Fig. 1. In light of this observation, we propose a novel LLE paradigm incorporating aesthetic assessment to effectively model human subjective preferences and improve both subjective experience and objective evaluation of the enhanced images.



(a) Normal images with an average Aesthetic Score 7.081 (b) Low-light images with an average Aesthetic Score 5.694

Figure 1: Sample images on DPChallenge protect[1]. Each image receives dozens to hundreds of user ratings, ranging from 1 to 10. Higher scores indicate higher aesthetic quality.

However, embedding aesthetic assessment into LLE is non-trivial, as aesthetics are highly subjective and personalized evaluations that vary between individuals. These personalized user preferences can bring confusion to model learning. In addition, the manual aesthetic retouching of photographs is a highly causal and progressive process, which is difficult to replicate using existing LLE methods. To tackle these challenges, we propose the following solutions. First, we introduce a well-trained 'aesthetic oracle network' to construct an Aesthetic Assessment Module that can produce general aesthetic preferences, reducing penalization and increasing the method's versatility and aesthetic appeal. Second, we apply reinforcement learning to interact with the environment (*i.e.*, the Aesthetic Assessment Module) to calculate rewards; we treat LLE as a Markov decision process, decomposing the augmented mapping relationship into a series of itera-

[1]Dong Liang is the corresponding author. We thank Prof. Jie Qin from NUAA and Dr. Tianyu Ding from Microsoft Seattle for their comments. This work is supported by the National Natural Science Foundation of China under grant 62272229, and the Natural Science Foundation of Jiangsu Province under grant BK20222012.

[2]https://www.dpchallenge.com/

tions through an Aesthetic Policy Generation Module, thus realizing progressive LLE adjustment. Third, we develop a group of complementary rewards, including aesthetics quality, feature preservation, and exposure control, to preserve better subjective visual experience and objective evaluation. All the rewards are non-reference, indicating that they do not require paired training images.

The main contributions of this paper are three-fold:

- We propose a new paradigm for LLE by integrating aesthetic assessment, leveraging aesthetic scores to mimic human subjective evaluations as a reward to guide LLE. To our knowledge, this is the first attempt to solve low-light image enhancement using aesthetics.
- We devise aesthetics-guided LLE (ALL-E) through reinforcement learning and treat LLE as a Markov decision process, which divides the LLE process into two phases: aesthetic policy generation and aesthetic assessment.
- Our method is evaluated against state-of-the-art competitors through comprehensive experiments in terms of visual quality, no- and full-referenced image quality assessment, and human subjective surveys. All results consistently demonstrate the superiority of ALL-E.

2 Related Work

Low-Light Image Enhancement

Traditional Approaches. Early efforts commonly presented heuristic priors with empirical observations to address the LLE problems [Pizer *et al.*, 1990; Land, 1977; Guo *et al.*, 2016]. Histogram equalization [Pizer *et al.*, 1990] used a cumulative distribution function to regularize the pixel values to achieve a uniform distribution of overall intensity levels in the image. The Retinex model [Land, 1977] and its multi-scale version [Jobson *et al.*, 1997] decomposed the brightness into illumination and reflections, which were then processed separately. Guo *et al.* [Guo *et al.*, 2016] introduced a structural prior before refining the initially obtained illumination map and synthesizing the enhanced image according to the Retinex theory. However, these constraints/priors must be more self-adaptive to recover image details and color, avoiding the washing out details, local under/over-saturation, uneven exposure, or halo artifacts.

Deep Learning Approaches. Lore *et al.* [Lore *et al.*, 2017] proposed a variant of the stacked sparse denoising autoencoder to enhance the degraded images. RetinexNet [Wei *et al.*, 2018; Zhang *et al.*, 2019] leveraged a deep architecture based on Retinex to enhance low-light images. RUAS [Liu *et al.*, 2021] constructed the overall LLE network architecture by unfolding its optimization process. EnlightenGAN [Jiang *et al.*, 2021] introduced GAN-based unsupervised training on unpaired normal-light images for LLE. Zero-DCE [Guo *et al.*, 2020] transformed the LLE task into an image-specific curve estimation problem. SCL-LLE [Liang *et al.*, 2022] cast the image enhancement task as multi-task contrast learning with unpaired positive and negative images and enabled interaction with scene semantics. However, all the above methods ignore the human subjective preferences of LLE.

Aesthetic Quality Assessment

Although there is no aesthetics-guided LLE method, and the focus of LLE is not on assessing the aesthetic quality of a given image, our work relates to this research domain in the sense that LLE aims at improving image quality.

Image aesthetics has become a widely researched topic in current computer vision. Image aesthetic quality assessment aims to simulate human cognition and perception of beauty to automatically predict how beautiful an image looks to a human observer. Previous attempts have trained convolutional neural networks for binary classification of image quality [Lu *et al.*, 2014; Ma *et al.*, 2017] or aesthetic score regression [Kong *et al.*, 2016]. Assessing visual aesthetics has practical applications in areas such as image retrieval [Yu and Moon, 2004], image recommendation [Yu *et al.*, 2021a], and color correction [Deng *et al.*, 2018].

Since the highly differentiated aesthetic preference, image aesthetics assessment can be divided into two categories: generic and personalized image aesthetics assessment (GIAA and PIAA)[Ren *et al.*, 2017]. [Ren *et al.*, 2017] addressed PIAA problem by leveraging GIAA knowledge on user-related data so that that model can capture aesthetic "offset". Later, research work attempted to learn PIAA from various perspectives, such as multi-modal collaborative learning [Wang *et al.*, 2018], meta-learning [Zhu *et al.*, 2020], multi-task learning [Li *et al.*, 2020] etc. Due to the high subjectivity of PIAA tasks, which pay more attention to the differences in personality factors, we aim to introduce "general aesthetic preferences"/generic image aesthetics assessment (GIAA) to LLE. To our knowledge, the proposed scheme in this paper is the first attempt to solve the LLE problem using aesthetics to preserve better both subjective visual experience and objective evaluation.

Reinforcement Learning for Image Processing

After deep Q-network achieved human-level performance on Atari games, there has been a surge of interest in deep reinforcement learning (DRL). For image-processing tasks, Yu *et al.* [Yu *et al.*, 2018] proposed RL-Restore to learn a policy of selecting appropriate tools from a predefined toolbox to restore the quality of corrupted images gradually. Park *et al.* [Park *et al.*, 2018] presented a DRL-based method for color enhancement and a distortion-recovery training scheme that only requires high-quality reference images for training.

While these methods focus on global image restoration, Furuta *et al.* [Furuta *et al.*, 2019] proposed pixelRL to enable pixel-wise image restoration, which extended DRL to pixel-level reinforcement learning, making it more flexible in dealing with image problems. Similarly, Zhang *et al.* [Zhang *et al.*, 2021a] proposed a novel DRL-based method for achieving LLE at the pixel level. In contrast, our DRL network learns with an image aesthetic reward to obtain LLE results that try to satisfy universal users.

3 Methodology

3.1 From Aesthetic Annotation to LLE

Many factors can affect the beauty of images, including the richness of color, correct exposure, depth of view, resolution,

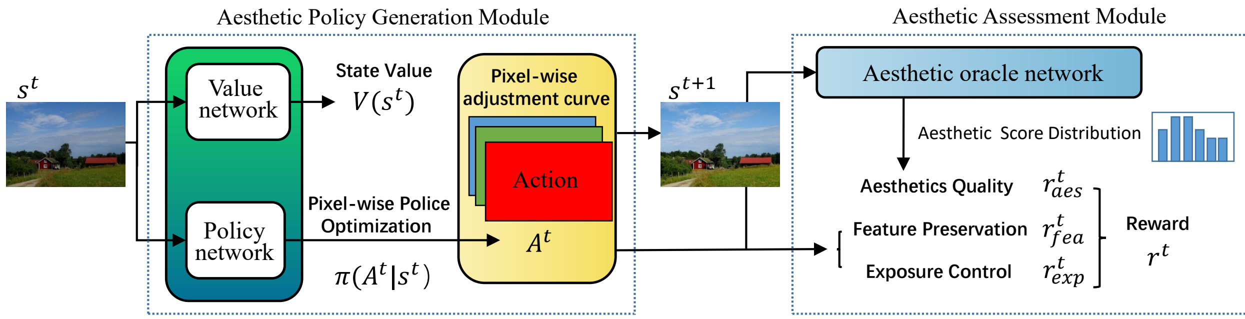


Figure 2: Overall architecture of the proposed ALL-E. It includes an aesthetic policy generation module and an aesthetic assessment module.

high-level semantics, and so on. Previous work [Lu *et al.*, 2014; Ma *et al.*, 2017; Kong *et al.*, 2016; Yu and Moon, 2004] have preliminarily analyzed these factors to generate the aesthetic rating of an image, facilitating manual aesthetic annotation for downstream tasks.

As aesthetics is a highly subjective and personalized evaluation with individual differences, [Kang *et al.*, 2020] modeled an aesthetic mean opinion score (AMOS) as a weighted sum of its four relatively independent attributes to generate more reliable aesthetic annotation,

$$AMOS \cong 0.288f_1 + 0.288f_2 + 0.082f_3 + 0.342f_4 \quad (1)$$

where f is the scores of attribute index $\{1, 2, 3, 4\}$ of an image. 1 is the light/color attribute, 2 is the composition and depth attribute, 3 is the imaging quality attribute, and 4 is the semantic attribute. To derive the specific importance (the weight values) in AMOS, [Kang *et al.*, 2020] directly elicited from more than 30 observers the importance of each attribute in forming their overall aesthetic opinion. The observers had to indicate which factor(s) influenced their overall aesthetic score among the rated attributes. From Eq.(1), we observe that light/color is one of the three most important attributes, contributing approximately 30%. Another work [Yang *et al.*, 2022] utilizes a Pearson correlation coefficient between the light condition of an image and its aesthetic rating, archiving a high correlation coefficient of 0.67.

Motivated by the above findings, we propose an aesthetics-guided LLE. The first step is to select an aesthetic 'oracle' that can provide general aesthetic preferences with versatility and reflect popular aesthetics. In our implementation, we employ a lightweight aesthetic network [Esfandarani and Milanfar, 2018] with a VGG-16 backbone as the aesthetic 'oracle', trained on the AVA dataset [Murray *et al.*, 2012]. Though NIMA [Esfandarani and Milanfar, 2018] (81.5%) stands at the middle level in terms of accuracy (DMA-Net [Lu *et al.*, 2015] 75.4%; MTCNN [Kao *et al.*, 2017] 79.1%; Pool-3FC [Hosu *et al.*, 2019] 81.7%; ReLIC [Zhao *et al.*, 2020] 82.35%), NIMA is with less computational complexity compared with other available models (ReLIC with a MobileNetV2 backbone has 2.7 times more parameters than NIMA with a MobileNetV1 backbone, while MobileNetV2 has fewer parameters than MobileNetV1). AVA is a database with 250,000 photos evaluated for aesthetic quality, with crowd-sourcing voting on the aesthetics for each image (from 78 to 549 people for each image). The aesthetic oracle network pre-trained on this vast dataset yields trustworthy preferences for broad aesthetics.

Since human aesthetic image retouching is a dynamically and explicitly progressive process that causally interacts with the current state of the image, we treat LLE as a Markov decision process, decomposing it into a series of iterations. To mimic this process, we design a reinforcement-learning-based Aesthetic Policy Generation Module, which interacts with the environment – Aesthetic Assessment Module to obtain rewards and provide optimized actions to realize progressive LLE adjustment. As shown in Fig. 2, our ALL-E consists of an Aesthetic Policy Generation Module based on an asynchronous advantage actor-critic network (A3C) [Mnih *et al.*, 2016] to generate the action A^t, and an Aesthetic Assessment Module based on an aesthetic oracle network [Esfandarani and Milanfar, 2018] and a series of loss function to generate the reward r^t. At the t-th step, given an image s^t, the Aesthetic Policy Generation Module generates an enhanced image s^{t+1} via A^t, which is then fed to the Aesthetic Assessment Module to generate r^t, then progressively complete image enhancement until n steps.

3.2 Aesthetic Policy Generation

The reason to use A3C [Mnih *et al.*, 2016] is that it reports pixel-wise action performance with efficient training. A3C is an actor-critic method consisting of two sub-networks: a value network and a policy network denoted as θ_v and θ_p, respectively. Both networks use the current state image s^t as the input at the t-th step. The value network outputs the value $V(s^t)$, which represents the expected total discounted rewards from state s^t to s^n, and indicates how good the current state is:

$$V(s^t) = \mathbb{E}\left[R^t \mid s^t\right] \quad (2)$$

where R^t is the total discounted reward:

$$R^t = \sum_{i=0}^{n-t-1} \gamma^i r^{t+i} + \gamma^{n-t} V(s^n) \quad (3)$$

where γ^i is the i-th power of the discount factor γ, r^t is the immediate reward (will be introduced in Section 3.4) at t-th step. As some actions can affect the reward after many steps by influencing the environment, reinforcement learning aims to maximize the total discounted reward R^t rather than the immediate reward r^t.

The gradient for θ_v is computed as follows:

$$d\theta_v = \nabla_{\theta_v}(R^t - V(s^t))^2 \tag{4}$$

The policy network outputs the probability of taking action $A^t \in A.S.$ (the action space, will be introduced in the following subsection) through softmax, denoted as $\pi(A^t|s^t)$. The output dimension of the policy network is equal to $|A|$.

To measure the rationality of selecting a specific action A^t in a state s^t, we define the advantage function as:

$$G(A^t, s^t) = R^t - V(s^t) \tag{5}$$

It directly gives the difference between the performance of action A^t and the mean value of the performance of all possible actions. If this difference (*i.e.* the advantage of the chosen action) is greater than 0, then it indicates that action A^t is better than the average and is a reasonable choice; if the difference is less than 0, then it implies that action A^t is inferior to the average and should not be selected.

The gradient for θ_p is computed as follows:

$$d\theta_p = -\nabla_{\theta_p} \log \pi(A^t|s^t) G(A^t, s^t) \tag{6}$$

3.3 Action Space Setting

Human experts often manually retouch photographs through curve adjustments in retouching software, where the curve parameters depend only on the input image. Usually, curves for challenging low-light images are of a high order. [Guo *et al.*, 2020] suggests that this procedure can be realized by recurrently applying the low-order curves. In this work, we apply a second-order pixel-wise adjustment curve (PAC) at each step t. We enhance an input image s^t by iteratively applying a PAC-based action $A^t(x)$ at the t-th step,

$$s^{t+1}(x) = s^t(x) + A^t(x)s^t(x)(1 - s^t(x)) \tag{7}$$

where x denotes the pixel coordinates. Eq.(7) models the enhancement of a low-light image as a sequential action-making problem by finding the optimal pixel-wise parameter map $A^t(x)$ for light adjustment curves at each step t. Therefore, our optimization goal is to find an optimal light adjustment action sequence. To achieve this, we need a metric (the reward) to measure the light aesthetic of an image s^t. The exact calculation is described in the following subsection.

Fundamentally, low-light image enhancement can be regarded as the search for a mapping function F, such that $s_H = F(s_L)$ is the desired image, which is enhanced from the input image s_L. In our design, the mapping function F is represented as the ultimate form of multiple submappings $\{A^1, A^2, \ldots, A^n\}$ continuously and iteratively augmented, where A^t is constrained in a predetermined range of the action space ($A.S.$). The range of $A.S.$ is crucial to the performance of our method, as too narrow a range will result in insufficient improvements, and too extensive a range will result in excessive search space.

Here, we empirically set the range of $A.S. \in [-0.5, 1]$ with a graduation interval $1/18$. This setting ensures that:

1) The brightness of each pixel is in the normalized range $[0, 1]$, according to Eq.(7);

2) As the brightness of some pixels may need to be reduced, a negative range $[-0.5, 0]$ is also set;

3) PAC is monotonous while also alleviating the cost of searching for suitable PAC for low-light image enhancement.

3.4 Reward Function

This section introduces three complementary rewards, including aesthetics quality, feature preservation, and exposure control, to preserve better subjective visual experience and objective evaluation. Note that all the rewards are non-reference, requiring no paired training images.

Aesthetics Quality Reward. As mentioned above, the aesthetic quality score of an image is closely correlated with several factors. In this work, we focus on dynamically adjusting and improving the brightness by the aesthetic score of an image. Therefore, utilizing the aesthetic score as a direct reward function would be an inadequate representation of the desired outcome. Instead, the difference in aesthetic scores between the original and enhanced images is employed as the reward for the currently selected action. The image aesthetics quality reward r^t_{aes} is:

$$r^t_{aes} = \sum_{k=1}^{K} k(P_k(s^{t+1}) - P_k(s^t)) \tag{8}$$

where K denotes the range of ratings for the aesthetic scores of the images, *i.e.*, [1, 10], and P denotes the probability of each rating. s^t denotes the state of the image at t-th step, s^{t+1} denotes the state of the image at $t+1$-th step.

Feature Preservation Reward. Since color naturalness is a concern in low-light image enhancement, we introduce a color constancy term incorporating an illumination smoothness penalty term as the feature preservation reward. It is based on the gray-world color constancy hypothesis [Buchsbaum, 1980; Guo *et al.*, 2020], which posits that the average pixel values of the three channels tend to be of the same value. r^t_{fea} constrains the ratio of the three channels to prevent potential color deviations in the enhanced image. In addition, to avoid aggressive and sharp changes between neighboring pixels, an illumination smoothness penalty term is also embedded in $r^t_{fea} =$

$$\sum_{\forall(p,q)\in\xi} (J^p - J^q)^2 + \lambda \frac{1}{n} \sum_{t=1}^{n} \sum_{p\in\xi} (\left|\nabla_x (A^t)^p\right| + \left|\nabla_y (A^t)^p\right|) \tag{9}$$

where $\xi = \{R, G, B\}$, J^p denotes the average intensity value of p channel in an image, (p, q) represents a pair of channels, n is the number of the steps, and ∇_x and ∇_y denote the horizontal and vertical gradient steps, respectively. We set λ to 100 in our experiments to achieve the best results.

Exposure Control Reward. The exposure control reward r^t_{exp}, which is a loss function widely used in the recent literature [Guo *et al.*, 2020; Zhang *et al.*, 2021a], measures the deviation of the average intensity value of a local region from a predefined well-exposedness level in RGB color space:

$$r^t_{exp} = \frac{1}{B} \sum_{b=1}^{B} |Y_b - E| \tag{10}$$

(a) Input (b) LIME (c) Retinex-Net (d) ISSR (e) Zero-DCE

(f) EnlightenGAN (g) RUAS (h) ReLLIE (i) SCL-LLE (j) Ours

Figure 3: Examples of enhancement results on LOL test dataset.

Methods	NIQE ↓	UNIQUE ↑	PSNR ↑	SSIM ↑	User study↓
Input	6.749	-0.144	7.773	0.194	4.333
(TIP'17) LIME	8.058	0.333	14.221	0.521	3.855
(BMVC'18) Retinex-Net	8.879	-0.026	16.774	0.424	3.277
(ACMMM'20) ISSR	3.872	0.739	12.469	0.525	3.950
(CVPR'20) Zero-DCE	7.767	0.335	14.860	0.562	3.286
(TIP'21) EnlightenGAN	5.807	0.546	17.654	0.666	3.156
(CVPR'21) RUAS	6.340	0.427	16.405	0.503	3.431
(ACMMM'21) ReLLIE	4.535	1.133	**19.454**	0.756	2.677
(AAAI'22) SCL-LLE	4.571	0.544	12.354	0.591	3.270
Ours	**3.774**	**1.227**	18.216	**0.763**	**2.450**

Table 1: NIQE ↓, UNIQUE ↑, PSNR ↑, SSIM ↑ and User study ↓ scores on LOL test dataset.

Methods	DICM		LIME		MEF		VV		NPE		Average	
	NIQE	UN.	NIQE	UN.	NIQE	UN.	NIQE	UN.	NIQE	UN.	NIQE	UN.
Input	4.26	0.72	4.36	0.70	4.26	0.72	3.52	**0.74**	4.32	**1.17**	4.13	0.75
(TIP'17) LIME	3.75	0.78	3.85	0.53	3.65	0.65	**2.54**	0.44	4.44	0.93	3.55	0.69
(BMVC'18) Retinex-Net	4.47	0.75	4.60	0.52	4.41	0.97	2.70	0.36	4.60	0.81	4.13	0.69
(ACMMM'20) ISSR	4.14	0.59	4.17	0.83	4.22	0.87	3.57	0.62	4.02	0.99	4.03	0.68
(CVPR'20) Zero-DCE	3.56	0.82	3.77	0.73	3.28	1.22	3.21	0.48	3.93	1.07	3.50	0.81
(TIP'21) EnlightenGAN	3.55	0.63	**3.70**	0.49	**3.16**	1.03	3.25	0.58	3.95	1.07	3.47	0.69
(CVPR'21) RUAS	5.21	-0.17	4.26	0.34	3.83	0.73	4.29	-0.04	5.53	0.13	4.78	0.04
(AAAI'22) SCL-LLE	3.51	0.87	3.78	0.76	3.31	1.25	3.16	0.49	3.88	1.08	3.46	0.85
(CVPR'22)Uretinex-net	3.95	0.85	4.34	**0.93**	3.79	1.18	3.01	0.51	4.69	0.99	3.83	0.84
(ICCV'21)Zhao *et al.*	3.68	**0.91**	4.16	0.79	3.83	0.97	3.01	0.57	**3.69**	1.06	3.61	0.84
(CVPR'22)Ma *et al.*	4.11	0.11	4.21	0.35	3.63	1.04	2.92	0.05	4.47	0.21	3.87	0.35
Ours	**3.49**	0.88	3.78	0.80	3.32	**1.27**	3.08	0.49	3.85	1.10	**3.45**	**0.88**

Table 2: NIQE ↓ and UNIQUE (UN.) ↑ scores on DICM, LIME, MEF, VV, and NPE datasets.

where B represents the number of non-overlapping local regions of size 16×16, Y_b is the average intensity value of a local region b in s^{t+1}. According to [Guo *et al.*, 2020; Zhang *et al.*, 2021a], E is set to 0.6. For a given enhanced image, the immediate reward r^t at a current state s^t is:

$$r^t = w_1 r^t_{aes} - w_2 r^t_{fea} - w_3 r^t_{exp} \tag{11}$$

where w_1, w_2 and w_3 are tunable hyperparameters. As introduced in Section 3.2, the goal of reinforcement learning is to maximize the total discounted reward R^t in Eq.(3) with this immediate reward r^t.

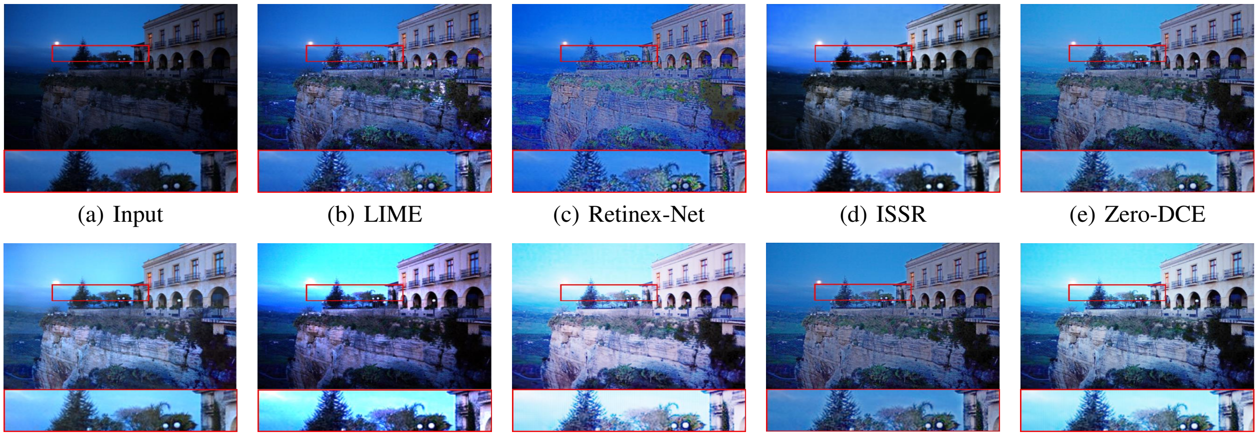
3.5 Efficient Training Details

We use 485 low-light images of the LOL dataset [Wei *et al.*, 2018] to train the proposed framework. We resize the training images to the size of 244×244. The maximum number of training epochs was set to 1000, with a batch size of 2. We train our framework end-to-end while fixing the weights of the aesthetic oracle network. Our framework is implemented in PyTorch on an NVIDIA 1080Ti GPU. The model is optimized using the Adam optimizer with a learning rate of $1e^{-4}$. The total number of steps in the training phase is set to $n = 6$. In accordance with the training phase, the total number of steps is also set to $n = 6$ in the testing phase. Under these settings, training 1000 epochs costs about one day.

4 Experiments

4.1 Benchmark Evaluations

We compare our method with several state-of-the-art methods: LIME [Guo *et al.*, 2016], Retinex-Net [Wei *et al.*, 2018], ISSR [Fan *et al.*, 2020], Zero-DCE [Guo *et al.*, 2020], EnlightenGAN [Jiang *et al.*, 2021], RUAS [Liu *et al.*, 2021], ReLLIE [Zhang *et al.*, 2021a], SCL-LLE [Liang *et al.*, 2022], Uretinex-net[Wu *et al.*, 2022], Zhao et al.[Zhao *et al.*, 2021], Ma et al.[Ma *et al.*, 2022]. The results of the above meth-



(a) Input (b) LIME (c) Retinex-Net (d) ISSR (e) Zero-DCE

(f) EnlightenGAN (g) RUAS (h) ReLLIE (i) SCL-LLE (j) Ours

Figure 4: Comparison of our method and the state-of-the-art methods over LIME dataset with zoom-in regions. Our method enables the enhanced images to look more realistic and recovers better details in both foreground and background.

ods are reproduced by the publicly available models provided with the recommended test settings. To thoroughly evaluate the proposed method, a comprehensive set of experiments were conducted, including a visual quality comparison, image quality assessment, and human subjective survey, which are discussed in the following sections.

Visual Quality Comparison

We present the visual comparisons on typical low-light images in LOL test dataset [Wei *et al.*, 2018] and LIME [Guo *et al.*, 2016] dataset. We first investigate whether the proposed method achieves visually pleasing results in terms of brightness, color, contrast, and naturalness. We observe from Fig. 3 and Fig. 4 that the images enhanced by our method are aesthetically acceptable and do not cause any discernible noise and artifacts. Specifically, LIME causes color artifacts at strong local edges; Retinex-Net and EnlightenGAN cause local color distortion and lack of detail; ISSR and RUAS produce severe global and local over/under-exposure; Zero-DCE and SCL-LLE under-enhance extremely dark images, while ReLLIE over-enhances.

Fig. 5 demonstrates the visualization of the enhancing procedure of the proposed method. As the enhancement steps t progress, the image's brightness increases. In our test experiments, step $t = 6$ yields the best visual performance, which is rational as the total number of steps is set to $n = 6$ in the training phase. When continuing to enhance the image with an additional step ($t = 7$), the image may tend to be over-enhanced, as shown in Fig. 5 (h). After many experiments, we found that $t = 6$ is the global best in most cases. in practice, the optimal enhancement step for a specific image may be different. The image sequences of all steps can be listed for people to choose their preferred one.

Image Quality Assessment (IQA)

For quantitative comparison, we use two non-reference evaluation indicators, Natural Image Quality Evaluator (NIQE) [Mittal *et al.*, 2013], and UNIQUE [Zhang *et al.*, 2021b]. NIQE is a well-known no-reference image quality assessment for evaluating image restoration without ground truth and providing quantitative comparisons. Since NIQE is regarded as poorly correlated with subjectivity, we also adopt UNIQUE, a metric for non-reference evaluation that is more rational and closer to subjective human opinion.

For full-reference image quality assessment, we employ the Peak Signal-to-Noise Ratio (PSNR, dB) and Structural Similarity (SSIM) metrics to compare the performance of various approaches quantitatively.

Table 1 and 2 summarizes the performances of our technique and thetechnique state-of-the-art methods on the test images of LOL test dataset [Wei *et al.*, 2018], DICM [Lee *et al.*, 2012], LIME [Guo *et al.*, 2016], MEF [Ma *et al.*, 2015], VV[1], and NPE [Wang *et al.*, 2013]. DICM, LIME, MEF, VV, NPE are ad hoc test datasets, including 64, 10, 8, 24, 17 images, respectively. They are widely used in LLE testing: SCL-LLE [Liang *et al.*, 2022], EnlightenGAN [Jiang *et al.*, 2021], Zero-DCE [Guo *et al.*, 2020] *et al.*. Images in them are diverse and representative: DICM is mainly landscaped with extreme darkness; LIME focuses on dark street landscapes; MEF focuses on dark indoor scenes and buildings; VV is mostly backlit and portraits; NPE mainly includes natural scenery in low light. Note that only the PSNR of our method is not ranked first. We believe this is because the aesthetic reward focuses more on the global aesthetic evaluation and is insufficiently responsive to local noise.

Human Subjective Survey

We conduct a human subjective survey (user study) for comparisons. Each image in the LOL test dataset was enhanced by nine methods (LIME, Retinex-Net, ISSR, Zero-DCE, EnlightenGAN, RUAS, ReLLIE, SCL-LLE, and our approach), 25 human volunteers were asked to rank the enhanced images. These subjects are instructed to consider: 1) Whether or not noticeable noise is present in the images; 2) If the images contain over or underexposure artifacts; 3) Whether or not the images display non-realistic color or texture distortion.

[1] https://sites.google.com/site/vonikakis/datasets

(a) Input (b) $t = 1$ (c) $t = 2$ (d) $t = 3$

(e) $t = 4$ (f) $t = 5$ (g) $t = 6$ (h) $t = 7$

Figure 5: An example of the proposed method with different enhancement steps.

(a) Input (b) w/o r^t_{aes} (c) Baseline $A.S.$ settings (d) Ours (e) Ground truth

Figure 6: Ablation study on the contribution of each component.

We assign each image a score between 1 and 5; the lower the value, the higher the image quality. The final results are shown in Table 1. We can see that our method achieves the highest score.

4.2 Ablation Study

To demonstrate the effectiveness of the aesthetic rewards and the action space configuration proposed by our technique, we performed several ablation experiments. Since r^t_{fea} and r^t_{exp} are demonstrated to be valid in the recent literature [Guo *et al.*, 2020; Zhang *et al.*, 2021a], we consider them to be the baseline rewards without ablation studies on them.

Action Space and Aesthetics Quality Reward. Regarding the the action space ($A.S.$) settings, the baseline follows [Zhang *et al.*, 2021a]: the range of $A.S. \in [-0.3, 1]$ with graduation as 0.05. In comparison, our settings fall within the range of $A.S. \in [-0.5, 1]$ with a graduation of $1/18$. Specifically, we design two experiments by removing the aesthetics quality reward component and keeping the baseline settings.

Subjective Experience. The visualization of the effects of action space and aesthetics quality reward r^t_{aes} are shown in Fig. 6. The absence of r^t_{aes} rendered the image gloomy and unappealing, while improper action space settings led to overexposure in certain portions of the enhanced image.

Objective Evaluation. Table 3 shows the NIQE, UNIQUE, PSNR, and SSIM scores under each experiment. Note that, from Table 3, the absence of r^t_{aes} does not appear to have a significant impact on the overall outcomes; however, the absence of a suitable action space ($A.S.$) setting appears to have a significant negative impact on performance. Since our experimental setup relies on the presence of different settings of the action space and image aesthetics quality reward, the absence of any component would result in subpar results. The aesthetics-guided action space setting can achieve the best results.

Our method can be seen as a compromise between aesthetics and Image Quality Assessment (IQA). In some specific scenarios, there would be situations where a high aesthetic score with poor IQA. The Aesthetics Quality Reward r_{aes} considers general aesthetics, while the Feature Preservation Reward r_{fea} and the Exposure Control Reward r_{exp} in our method ensure relatively good IQA.

Our $A.S.$	r^t_{aes}	NIQE ↓	UNIQUE ↑	PSNR ↑	SSIM ↑
		4.822	1.131	13.920	0.697
✓		4.766	1.221	15.211	0.723
	✓	4.131	0.912	14.526	0.682
✓	✓	**3.774**	**1.227**	**18.216**	**0.763**

Table 3: Ablation study. NIQE ↓, UNIQUE ↑, PSNR ↑ and SSIM ↑ scores on LOL test dataset.

5 Conclusion

We have proposed an effective aesthetics-guided reinforcement learning method to solve the LLE problem. ALL-E illustrates how we can leverage aesthetics to balance the subjective and the output. Unlike most existing learning-based methods, using the aesthetic policy generation and aesthetic assessment modules, our method treats LLE as a Markov decision process to realize progressive learning. With aesthetic assessment scores as a reward, general human subjective preferences are introduced, which aids in producing aesthetically pleasing effects, *i.e.* it interacts with the environment (aesthetic assessment module) to yield a reward to mimic a human's image retouching process.

References

[Buchsbaum, 1980] Gershon Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin institute*, 1980.

[Cho *et al.*, 2020] Se Woon Cho, Na Rae Baek, Ja Hyung Koo, Muhammad Arsalan, and Kang Ryoung Park. Semantic segmentation with low light images by modified cyclegan-based image enhancement. *IEEE Access*, 8:93561–93585, 2020.

[Deng *et al.*, 2018] Yubin Deng, Chen Change Loy, and Xiaoou Tang. Aesthetic-driven image enhancement by adversarial learning. In *ACM International Conference on Multimedia*, pages 870–878, 2018.

[Esfandarani and Milanfar, 2018] Hossein Talebi Esfandarani and Peyman Milanfar. Nima: Neural image assessment. *IEEE Transactions on Image Processing*, 27(8):3998–4011, 2018.

[Fan *et al.*, 2020] Minhao Fan, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Integrating semantic segmentation and retinex model for low-light image enhancement. In *ACM International Conference on Multimedia, Virtual Event*, pages 2317–2325, 2020.

[Furuta *et al.*, 2019] Ryosuke Furuta, Naoto Inoue, and Toshihiko Yamasaki. Fully convolutional network with multi-step reinforcement learning for image processing. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 3598–3605, 2019.

[Guo *et al.*, 2016] Xiaojie Guo, Yu Li, and Haibin Ling. Lime: Low-light image enhancement via illumination map estimation. *IEEE Transactions on Image Processing*, pages 982–993, 2016.

[Guo *et al.*, 2020] Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In *CVPR*, pages 1780–1789, 2020.

[Hosu *et al.*, 2019] Vlad Hosu, Bastian Goldlucke, and Dietmar Saupe. Effective aesthetics prediction with multi-level spatially pooled features. In *CVPR*, pages 9375–9383, 2019.

[Jiang *et al.*, 2021] Yifan Jiang, Xinyu Gong, Ding Liu, Yu Cheng, Chen Fang, Xiaohui Shen, Jianchao Yang, Pan Zhou, and Zhangyang Wang. Enlightengan: Deep light enhancement without paired supervision. *IEEE Transactions on Image Processing*, pages 2340–2349, 2021.

[Jobson *et al.*, 1997] Jobson, Daniel, J., Rahman, Zia-ur, Woodell, Glenn, and A. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image Processing*, 1997.

[Kang *et al.*, 2020] Chen Kang, Giuseppe Valenzise, and Frédéric Dufaux. Eva: An explainable visual aesthetics dataset. In *Joint Workshop on Aesthetic and Technical Quality Assessment of Multimedia and Media Analytics for Societal Trends*, pages 5–13, 2020.

[Kao *et al.*, 2017] Yueying Kao, Ran He, and Kaiqi Huang. Deep aesthetic quality assessment with semantic information. *IEEE Transactions on Image Processing*, 26(3):1482–1495, 2017.

[Kong *et al.*, 2016] Shu Kong, Xiaohui Shen, Zhe Lin, Radomir Mech, and Charless Fowlkes. Photo aesthetics ranking network with attributes and content adaptation. In *European Conference on Computer Vision*, pages 662–679. Springer, 2016.

[Land, 1977] Edwin H Land. The retinex theory of color vision. *Scientific american*, pages 108–129, 1977.

[Lee *et al.*, 2012] Chulwoo Lee, Chul Lee, and Chang-Su Kim. Contrast enhancement based on layered difference representation. In *IEEE International Conference on Image Processing*, pages 965–968, 2012.

[Li *et al.*, 2020] Leida Li, Hancheng Zhu, Sicheng Zhao, Guiguang Ding, and Weisi Lin. Personality-assisted multi-task learning for generic and personalized image aesthetics assessment. *IEEE Transactions on Image Processing*, page 3898–3910, Feb 2020.

[Liang *et al.*, 2022] Dong Liang, Ling Li, Mingqiang Wei, Shuo Yang, Liyan Zhang, Wenhan Yang, Yun Du, and Huiyu Zhou. Semantically contrastive learning for low-light image enhancement. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 1555–1563, 2022.

[Liu *et al.*, 2021] Risheng Liu, Long Ma, Jiaao Zhang, Xin Fan, and Zhongxuan Luo. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. In *CVPR*, pages 10561–10570, 2021.

[Lore *et al.*, 2017] Kin Gwn Lore, Adedotun Akintayo, and Soumik Sarkar. Llnet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognition*, pages 650–662, 2017.

[Lu *et al.*, 2014] Xin Lu, Zhe Lin, Hailin Jin, Jianchao Yang, and James Z Wang. Rapid: Rating pictorial aesthetics using deep learning. In *ACM International Conference on Multimedia*, pages 457–466, 2014.

[Lu *et al.*, 2015] Xin Lu, Zhe Lin, Xiaohui Shen, Radomir Mech, and James Z Wang. Deep multi-patch aggregation network for image style, aesthetics, and quality estimation. In *Proceedings of the IEEE international conference on computer vision*, pages 990–998, 2015.

[Ma *et al.*, 2015] Kede Ma, Kai Zeng, and Zhou Wang. Perceptual quality assessment for multi-exposure image fusion. *IEEE Transactions on Image Processing*, pages 3345–3356, 2015.

[Ma *et al.*, 2017] Shuang Ma, Jing Liu, and Chang-Wen Chen. A-lamp: Adaptive layout-aware multi-patch deep convolutional neural network for photo aesthetic assessment. In *CVPR*, pages 722–731, 07 2017.

[Ma *et al.*, 2022] Long Ma, Tengyu Ma, Risheng Liu, Xin Fan, and Zhongxuan Luo. Toward fast, flexible, and robust low-light image enhancement. In *CVPR*, pages 5637–5646, 2022.

[Mittal *et al.*, 2013] Anish Mittal, R. Soundararajan, and A. Bovik. Making a "completely blind" image quality analyzer. *IEEE Signal Processing Letters*, pages 209–212, 2013.

[Mnih *et al.*, 2016] Volodymyr Mnih, Adria Puigdomenech Badia, Mehdi Mirza, Alex Graves, Timothy Lillicrap, Tim Harley, David Silver, and Koray Kavukcuoglu. Asynchronous methods for deep reinforcement learning. In *International Conference on Machine Learning*, pages 1928–1937. PMLR, 2016.

[Murray *et al.*, 2012] Naila Murray, Luca Marchesotti, and Florent Perronnin. Ava: A large-scale database for aesthetic visual analysis. In *CVPR*, pages 2408–2415. IEEE, 2012.

[Park *et al.*, 2018] Jongchan Park, Joon-Young Lee, Donggeun Yoo, and In So Kweon. Distort-and-recover: Color enhancement using deep reinforcement learning. In *CVPR*, pages 5928–5936, 2018.

[Pizer *et al.*, 1990] S. Pizer, R. Johnston, J. P. Ericksen, B. Yankaskas, and K. Muller. Contrast-limited adaptive histogram equalization: speed and effectiveness. *Conference on Visualization in Biomedical Computing*, pages 337–345, 1990.

[Ren *et al.*, 2017] Jian Ren, Xiaohui Shen, Zhe Lin, Radomir Mech, and David J. Foran. Personalized image aesthetics. In *2017 IEEE International Conference on Computer Vision (ICCV)*, Dec 2017.

[Ren *et al.*, 2019] Wenqi Ren, Sifei Liu, Lin Ma, Qianqian Xu, Xiangyu Xu, Xiaochun Cao, Junping Du, and Ming-Hsuan Yang. Low-light image enhancement via a deep hybrid network. *IEEE Transactions on Image Processing*, pages 4364–4375, 2019.

[Wang *et al.*, 2013] Shuhang Wang, Jin Zheng, Hai-Miao Hu, and Bo Li. Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE Transactions on Image Processing*, pages 3538–3548, 2013.

[Wang *et al.*, 2018] Guolong Wang, Junchi Yan, and Zheng Qin. Collaborative and attentive learning for personalized image aesthetic assessment. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, Jul 2018.

[Wei *et al.*, 2018] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. In *British Machine Vision Conference*, 2018.

[Wu *et al.*, 2022] Wenhui Wu, Jian Weng, Pingping Zhang, Xu Wang, Wenhan Yang, and Jianmin Jiang. Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement. In *CVPR*, pages 5901–5910, 2022.

[Xu *et al.*, 2020] Ke Xu, Xin Yang, Baocai Yin, and Rynson W.H. Lau. Learning to restore low-light images via decomposition-and-enhancement. In *CVPR*, 2020.

[Yang *et al.*, 2022] Yuzhe Yang, Liwu Xu, Leida Li, Nan Qie, Yaqian Li, Peng Zhang, and Yandong Guo. Personalized image aesthetics assessment with rich attributes. In *CVPR*, pages 19861–19869, 2022.

[Yu and Moon, 2004] So-Young Yu and Sung-Been Moon. An exploratory study of image retrieval using aesthetic impressions. *Journal of the Korean Society for information Management*, 21(4):187–208, 2004.

[Yu *et al.*, 2018] Ke Yu, Chao Dong, Liang Lin, and Chen Change Loy. Crafting a toolchain for image restoration by deep reinforcement learning. In *CVPR*, pages 2443–2452, 2018.

[Yu *et al.*, 2021a] Jing Yu, Lulu Zhao, et al. A novel deep cnn method based on aesthetic rule for user preferential images recommendation. *Journal of Applied Science and Engineering*, 24(1):49–55, 2021.

[Yu *et al.*, 2021b] Jun Yu, Xinlong Hao, and Peng He. Single-stage face detection under extremely low-light conditions. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, pages 3523–3532, October 2021.

[Zhang *et al.*, 2019] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. Kindling the darkness: A practical low-light image enhancer. In *ACM International Conference on Multimedia*, page 1632–1640, 2019.

[Zhang *et al.*, 2021a] Rongkai Zhang, Lanqing Guo, Siyu Huang, and Bihan Wen. Rellie: Deep reinforcement learning for customized low-light image enhancement. In *ACM International Conference on Multimedia*, pages 2429–2437, 2021.

[Zhang *et al.*, 2021b] Weixia Zhang, Kede Ma, Guangtao Zhai, and Xiaokang Yang. Uncertainty-aware blind image quality assessment in the laboratory and wild. *IEEE Transactions on Image Processing*, pages 3474–3486, 2021.

[Zhao *et al.*, 2020] Lin Zhao, Meimei Shang, Fei Gao, Rongsheng Li, Fei Huang, and Jun Yu. Representation learning of image composition for aesthetic prediction. *Computer Vision and Image Understanding*, 199:103024, 2020.

[Zhao *et al.*, 2021] Lin Zhao, Shao-Ping Lu, Tao Chen, Zhenglu Yang, and Ariel Shamir. Deep symmetric network for underexposed image enhancement with recurrent attentional learning. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 12075–12084, 2021.

[Zhu *et al.*, 2020] Hancheng Zhu, Leida Li, Jinjian Wu, Sicheng Zhao, Guiguang Ding, and Guangming Shi. Personalized image aesthetics assessment via meta-learning with bilevel gradient optimization. *IEEE Transactions on Cybernetics*, 52(3):1798–1811, Jun 2020.