

Article

Dark Light Image-Enhancement Method Based on Multiple Self-Encoding Prior Collaborative Constraints

Lei Guan ^{1,2}, Jiawei Dong ^{1,2}, Qianxi Li ^{1,2}, Jijiang Huang ^{1,*}, Weining Chen ¹ and Hao Wang ^{1,*}

¹ Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China; guanlei21@mailsucas.ac.cn (L.G.); dongjiawei22@mailsucas.ac.cn (J.D.); liqianxi22@mailsucas.ac.cn (Q.L.)

² University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: huangjijiang@opt.ac.cn (J.H.); wanghao@opt.ac.cn (H.W.)

Abstract: The purpose of dark image enhancement is to restore dark images to visual images under normal lighting conditions. Due to the ill-posedness of the enhancement process, previous enhancement algorithms often have overexposure, underexposure, noise increases and artifacts when dealing with complex and changeable images, and the robustness is poor. This article proposes a new enhancement approach consisting in constructing a dim light enhancement network with more robustness and rich detail features through the collaborative constraint of multiple self-coding priors (CCMP). Specifically, our model consists of two prior modules and an enhancement module. The former learns the feature distribution of the dark light image under normal exposure as an a priori term of the enhancement process through multiple specific autoencoders, implicitly measures the enhancement quality and drives the network to approach the truth value. The latter fits the curve mapping of the enhancement process as a fidelity term to restore global illumination and local details. Through experiments, we concluded that the new method proposed in this article can achieve more excellent quantitative and qualitative results, improve detail contrast, reduce artifacts and noise, and is suitable for dark light enhancement in multiple scenes.

Keywords: dark light enhancement; self-encoding prior; fidelity term; collaborative constraint



Citation: Guan, L.; Dong, J.; Li, Q.; Huang, J.; Chen, W.; Wang, H. Dark Light Image-Enhancement Method Based on Multiple Self-Encoding Prior Collaborative Constraints. *Photonics* **2024**, *11*, 190. <https://doi.org/10.3390/photonics11020190>

Received: 20 December 2023

Revised: 8 February 2024

Accepted: 11 February 2024

Published: 19 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Currently, digital image processing technologies centered around network frameworks are gradually shining brightly in various aspects of human life, such as unmanned driving [1], industrial production [2], video surveillance [3], military applications, remote sensing monitoring [4] and other fields. However, the application premise of these technologies is to obtain accurate and clear digital images for recognition and judgment, so as to carry out a series of downstream tasks of computer vision. However, there are often some uncontrollable factors in the process of image acquisition, and it is impossible to obtain images with good imaging quality, especially in the case of poor lighting conditions, such as cloudy, indoor, night, deep-space, underground and other working environments. The images obtained by the device may exhibit severe distortion and noise. The image gray-scale range is narrow and the contrast is low [5], which makes the human eye unable to obtain effective information. It also affects the recognition and processing of images by computer vision.

Due to the hardware iteration of computers and the construction of large-scale datasets, deep learning networks can learn and parameter-fit large-scale datasets through deeper and more complex network structures [6], but dark light enhancement is different from other image processing tasks. There is difficulty in obtaining authentic and effective labeled image datasets in the dark light environment, so it is difficult to fit the feature mapping based on the paired (normal and dark image) data. Although there are many algorithms that can perform unsupervised enhancement [7] based on prior knowledge, such as image

exposure control, color constancy, and color consistency, there are still some limitations to the enhancement effect. Second, the definition range of dark light environment is relatively wide. If the images obtained under different dark light environments and different light intensities are enhanced by pre-trained models, there will inevitably be overexposure, underexposure, blurring and other phenomena.

In various aspects, such as denoising, deblurring and brightness adjustment, efforts are made to improve image quality and enhance low-light images. This article proposes a collaborative constraint dark light enhancement algorithm based on multiple self-coding priors (CCMP). Compared with most self-supervised algorithms that learn the mapping of truth values to the original image, this algorithm cleverly utilizes multiple priors to obtain the essential features of the image, reduces the dependence on the dataset, improves the robustness of the algorithm, and adapts to more complex dark environments. Compared with unsupervised algorithms that rely solely on multiple loss functions for enhancement, this algorithm guides the direction of enhancement through prior terms, constrains image enhancement changes through fidelity terms, and balances the feature relationships of color, texture, saturation, etc. in the enhanced image, achieving better visual effects.

The complete enhancement process is shown in Figure 1. First, the deep nonlinear features of the image are mined through specific multiple self-coding network structures as prior constraints for subsequent enhancement networks. Second, in the enhanced network module, the global detailed color features of the image are extracted, and a new image loss evaluation function is designed as a fidelity term to balance the multi-dimensional features of the image. Finally, a collaborative constraint approach is used to achieve both subjective and objective optimization. The main contributions of this paper are as follows:

1. By using an autoencoder based on an LBP (local binary pattern) to learn the detailed texture features of dark light images, the interference of brightness color information in the enhancement process is suppressed.
2. The mask self-encoder based on the MCMC(Markov chain Monte Carlo) algorithm is used to effectively capture the important features in the data. The unsupervised feature learning method improves the robustness and adaptability of the enhancement process, and effectively filters the noise while reconstructing the image.
3. The image difference evaluation function is designed as the loss function of the data and multiple autoencoder networks are combined in the enhancement network as a priori terms to constrain the enhancement process, and the losses based on image structural analysis and image difference are combined to guide the enhancement process of the dark light image to improve the enhancement effect and robustness.

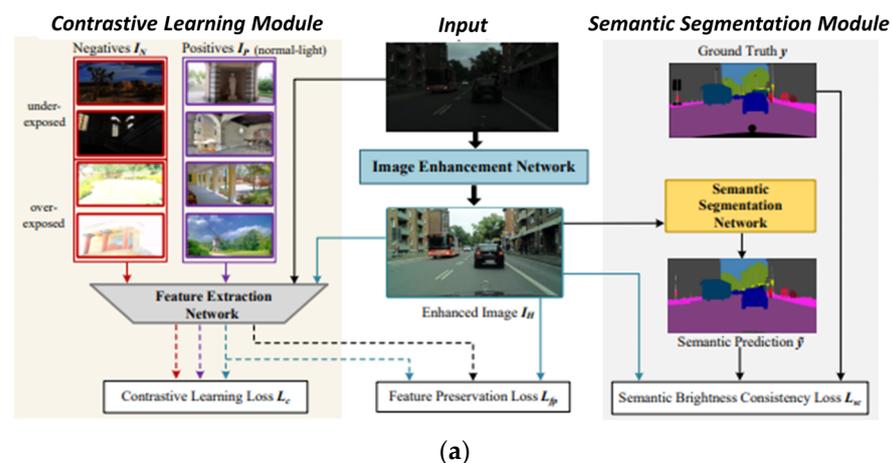


Figure 1. Cont.

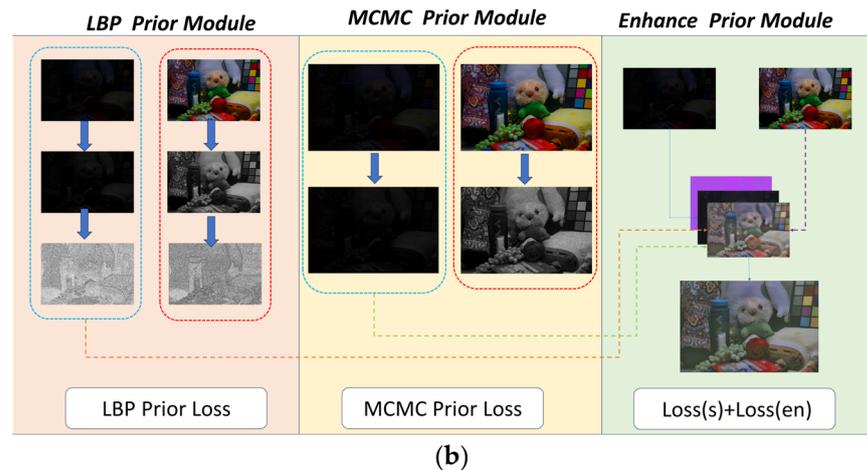


Figure 1. (a) This is a method of processing low-light images through semantic contrastive learning, which is divided into three parts: enhancement, contrastive learning and semantic segmentation. (b) This is a dark light enhancement network based on multiple prior collaborative constraints newly designed in this article. The blue box is used to enhance the network data flow, and the red box is not input into the network, but rather is only used to display the data processing method. Through the prior acquisition of LBP manipulation and MCMC mask prior, the information is obtained, and the differentiated loss is designed to constrain and enhance the loss of multiple pieces of prior information. The smooth loss acts on the enhanced network together to achieve a better image enhancement effect.

2. Related work

Up to now, the mainstream algorithms used for processing low-light images can be divided into three categories: traditional methods for grayscale and contrast, methods based on Retinex theory, and various variant network-enhancement frameworks based on deep learning.

- Traditional enhancement methods:

The method based on grayscale transformation is to map and optimize the grayscale values of each point in the image through fixed formulas and adjustment coefficients. The specific methods are piecewise, linear, nonlinear logarithmic function and gamma transformation. However, no matter what kind of solution, parameter optimization can only be based on experience or a large number of manual participations; it lacks an adaptive mechanism, and in the process of image enhancement, the overall gray distribution is not considered, which makes it easy to lose details and causes limited enhancement ability.

The dark light enhancement method based on Retinex theory is based on the Retinex theory proposed by Land [8]. To enhance the color consistency of the image, it is first decomposed before fusing it to enhance the low-light image. The algorithm based on this principle has been continuously improved. From the single-scale Retinex algorithm [9], to the multi-scale [10–12] weighted average MSR algorithm, and then to the color-recovery multi-scale MSRCR algorithm [13,14], some of the algorithm improvements have achieved breakthroughs in local and global dynamic compression, the enhancement of regional contrast, and the adjustment of color distortion. However, based on this theory, various methods still have problems such as a long processing time, limited model capacity, the inadaptability of hand-made constraints, the inability to process in batches quickly, glow appearing in shadow transitions in high-dynamic images and inferior color retention. Therefore, the results may produce strong noise, inappropriate exposure, insufficient details and unsaturated colors.

- Deep-learning methods:

In the enhanced networks, CNN is the most widely used framework, and most of them use multi-feature complementary fusion to optimize images. For example, the CNN network is used to settle the blind enhancement problem by using two-branch exposure fusion [15], the LLCNN of enhanced images is generated by using multi-level features [16], and the GLADNet of low-light images with incomplete global illumination is solved by using illumination prediction and original input-reconstruction details [17]. Jiang et al. [18] proposed the R2RNet low-illumination network, and Li et al. [19] proposed a dual-attention mechanism model for extracting local features to upgrade image quality.

At the same time, it is also a common idea to apply the Retinex algorithm to deep learning [20–22]. Liang et al. [23] used the Retinex network model to learn the mapping curve between low-light images and normal images. Liu [24] proposed a lightweight low-illumination RUAS model constructed using Retinex rules, which does not require any pairwise or non-paired supervision during the search process, and is lighter and more flexible than the previous architecture.

In addition, there are some enhancement methods for dark light images. Ren et al. [25] combined image perception and generative adversarial networks in the network, and the generated image visual quality was better than that achieved with an algorithm of the same level. Fan et al. [26] proposed a new deep network-integrating semantic segmentation for low-light image enhancement. Using semantic prior and signal structures, the illumination distribution and moderate noise were successfully processed, and good visual quality was achieved. Liu [27] designed a new event and image fusion transform (EIFT) module for event and image fusion, and a guided dual-branch (EGDB) module for weak light enhancement. Wang et al. [28] designed two independent networks to learn the characteristics of lighting and noise in images, which are interrelated to play a role in image optimization.

In summary, building various network models for optimizing low-light images is the mainstream approach. However, mainstream deep-learning methods rely on large datasets, the models have massive parameters, and they cannot achieve real-time enhancement requirements. Moreover, most of the enhancement models are only used for specific scenarios, only to adjust the image brightness, and lack the ability to denoise and deblur. Therefore, it is necessary to continuously optimize and improve the enhancement methods for low-light images.

- Prior learning based on mask auto-encoder:

When the traditional autoencoder learns the prior information of the image, because the compression and decompression process cannot accurately reconstruct the details and complexity of the image, it is unmanageable to deal with the complex image structure, and it is prone to overfitting the potential distribution characteristics of lost images. At the same time, the autoencoder is usually a learning method based on local features, which decodes and reconstructs according to the low-dimensional representation of the code. This may lead to poor performance of autoencoders in learning global structural information and an inability to capture the overall layout and semantic features of the target data. Kaiming He et al. [29] first proposed a mask autoencoder network in 2021. The asymmetric encoder and decoder design and a self-supervised pre-training model with a higher mask rate (75%) are used to mask the original image data. The image region of interest can be focused, so that the model can better capture key image features and structures.

Since then, the mask-based autoencoder network has shown impressive results in tasks such as noise reduction and super-resolution reconstruction, and many improved variants have been derived. CAE [30] improves two aspects of better representation learning; one is to separate the ‘representation learning’ and ‘pretext task’ as far as possible, and the other is to use a representation space to predict the masked patch. It utilizes the principle of perceptual similarity to improve image transfer performance [31]. Paper [32] used the masked image to directly return the HOG feature in the feature space to achieve network training. It does not depend on data augmentation, nor does it require tokenization, and

more importantly, it can be extended to large models. MSN [33] matches the representation of an image view containing a random mask with the unobstructed original image. On the basis of the Siamese Network, the mask patch strategy is used and prototype supervision is added. However, how to determine the optimal mask strategy, mask size and mask position is still a problem that needs further study.

3. Materials and Methods

As shown in Figure 2, the CCMP algorithm uses the U-net network combined with a mixed-attention mechanism. The data feature captured by the dual-specific autoencoder network is used as a prior term, and then the regularization prior term is used to constrain the enhancement process of the guided dark-light image.

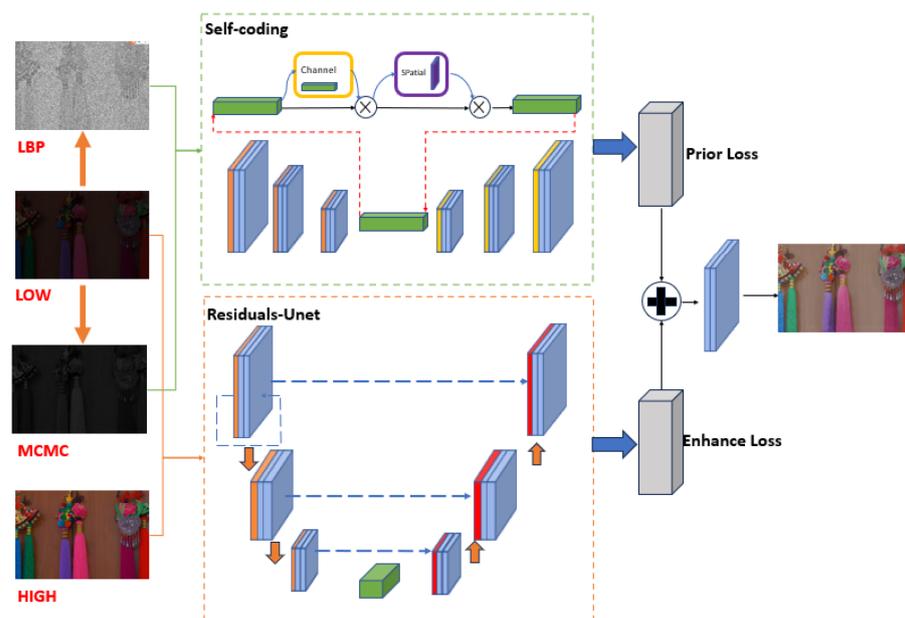


Figure 2. Enhanced network with multiple prior coordination constraints.

3.1. Self-Encoding Prior Based on Image LBP Processing

The prior term [34,35] is part of the constraint on the model based on prior knowledge or prior assumptions. By introducing an a priori term, the model can be regularized when data are insufficient, reducing the risk of overfitting, and guiding the model to learn in a more reasonable and reliable direction. Therefore, the information obtained by the prior term will directly affect the subsequent dark light enhancement process.

The current dark light enhancement network mostly processes the dark light image directly. In order to accurately obtain the image texture, detail and contour information, and suppress the influence of color and brightness in the original dark light image on the enhancement process, this paper obtains the prior process. By using a local binary pattern [36], the texture details of the image can be obtained by comparing the pixel values of adjacent small areas. The LBP process is as follows:

First, transforming an image from color to gray can effectively avert the influence of color. For the obtained gray image, $I(x, y)$ is used to represent the pixel value at the coordinate (x, y) in the image, and the target point grayscale value $c = I(x, y)$ is compared with the surrounding k pixels. The formula is as follows:

$$P_k = \begin{cases} 1, & \text{if } I(x_k, y_k) \geq c \\ 0, & \text{if } I(x_k, y_k) < c \end{cases} \quad (1)$$

The calculated P_k is arranged in order to obtain the LBP value as follows:

$$LBP(x, y) = P_7P_6P_5P_4P_3P_2P_1P_0. \tag{2}$$

The histogram is constructed by calculating the LBP value of all pixels, and the feature vector is obtained by normalization. The feature vector with local invariance and robustness to gray level change, illumination change and noise is obtained.

3.2. Mask Autoencoder Prior Based on Markov Monte Carlo Method

In this paper, a masked autoencoder based on the Markov chain Monte Carlo (MCMC) method [37,38] is proposed. The essence of the Monte Carlo method is to use randomness to solve the problem of certainty. Monte Carlo sampling utilizes the probability density function $f(x)$ of a known distribution to generate a sample x that obeys this distribution. Because the data are discrete, it is the same for color image data. We assume that x is uniformly distributed between $[a, b]$ and $p(x, i) = 1/(b - a)$. By introducing the Monte Carlo integral formula with probability distribution, we can obtain the following:

$$\int_a^b f(x)dx = \frac{1}{n} \sum_{i=0}^{n-1} \frac{f(x_i)}{1/(b - a)} = \frac{b - a}{n} \sum_{i=0}^{n-1} f(x_i). \tag{3}$$

As the probability distribution is not a common distribution, generally accept-reject sampling is used to obtain the distribution of the sample. However, using the Monte Carlo method, it is difficult to simulate high-dimensional distribution. The dataset in this paper contains images collected in a dark light environment, and more potential information is hidden in a high-dimensional space. Therefore, this paper solves the defects of the Monte Carlo algorithm (MCMC) in high-dimensional data of dark light images by using the Markov chain.

In the Markov chain [39] model state transition matrix, no matter from which initial probability distribution, substituted into the state transition matrix, the final state probability distribution tends to the same stable probability distribution, so the Markov chain sampling can obtain the corresponding sample set, which conforms to the stationary distribution, and then the Monte Carlo simulation is summed. Suppose that after n rounds, the Markov chain converges to our stationary distribution 1 as follows:

$$\pi_n(x) = \pi_{n+1}(x) = \pi_{n+2}(x) = \dots = \pi(x). \tag{4}$$

For each distribution $\pi_i(x)$,

$$\pi_i(x) = \pi_{i-1}(x)P = \pi_{i-2}(x)P^2 = \dots = \pi(x)P^i. \tag{5}$$

After n times, the sampling set conforms to the sample set corresponding to the stationary distribution, and then the Monte Carlo simulation is summed.

In this paper, the process of the Markov Monte Carlo method mask for dark light images is as follows. First, the mask M is initialized, and the energy function is defined, where $E(I, M)$ represents the energy function in image I , which is used to measure the difference between the image and the mask. The energy function formula is as follows:

$$E(I, M) = \sum |I - I \times M|, \tag{6}$$

Among them, I is the original image, M is equal to the mask, and $I \cdot M$ is the obtained mask graph. The symbol \sum represents the sum of all elements.

When designing the transition probability function, $T(M \rightarrow M')$ represents the transition probability from the current mask state M to the new mask state M' .

$$T(M \rightarrow M') = \min(1, e^{(E_i(M) - E_i(M'))}). \tag{7}$$

where M denotes the current mask state, M' denotes the new mask state, and the current mask energy and future mask energy are $E_i(M)$ and $E_i(M')$, respectively. The transfer probability function obtains changes in energy before and after flipping, and the difference is indexed to obtain a transition probability. According to the energy difference, we chose the smaller value between the minimum value and 1.0 as the transition probability to ensure that the transition probability range is within [0,1].

In N iterations, the position j of the mask is randomly selected to calculate the current energy: $E_j(M) = E(I, M)$. By flipping the mask position j , a new mask state M' is obtained. Then, the energy after flipping is calculated as $E_j(M') = E(I, M')$, the transition probability is calculated, and the flip operation is accepted as follows:

$$\alpha = \min(1, e^{(E_j(M')-E_j(M))}). \tag{8}$$

$$if \begin{cases} r > \alpha, M = M \\ r < \alpha, M = M' \end{cases} . \tag{9}$$

3.3. Loss Function

3.3.1. Loss of Image Local Contrast Difference

The details, textures and potential features of the original images learned by multiple specific autoencoder networks guide the image enhancement as the first validation and fidelity items in the enhancement network. For the prior image information, gray processing is performed first:

$$G_{(i,j)} = 0.2989 \times I_{(i,j)}^R + 0.5870 \times I_{(i,j)}^G + 0.1140 \times I_{(i,j)}^B, \tag{10}$$

where $I_{(i,j)}^R, I_{(i,j)}^G, I_{(i,j)}^B$ are the brightness values of the corresponding coordinate points in the red, green and blue channels of the image, respectively. The gradients in the X and Y directions are respectively calculated as follows:

$$R_{(i,j)}^X = [G_{(i-1, j+1)} + 2G_{(i, j+1)} + G_{(i+1, j+1)}] - [G_{(i-1, j-1)} + 2G_{(i, j-1)} + G_{(i+1, j-1)}], \tag{11}$$

$$R_{(i,j)}^Y = [G_{(i-1, j-1)} + 2G_{(i-1, j)} + G_{(i-1, j+1)}] - [G_{(i+1, j-1)} + 2G_{(i+1, j)} + G_{(i+1, j+1)}]. \tag{12}$$

The image difference measure is calculated as follows:

$$\downarrow_D = \frac{1}{W_{(i,j)}H_{(i,j)}} \sum_{i=1}^{W_{(i,j)}} \sum_{j=1}^{H_{(i,j)}} \sqrt{\{\sqrt{(R_{en}^X)^2 + (R_{en}^Y)^2} - \sqrt{(R_{gt}^X)^2 + (R_{gt}^Y)^2}\}^2}, \tag{13}$$

where $G_{(i,j)}$ is a grayscale image, and $R_{(i,j)}^X$ and $R_{(i,j)}^Y$ denote the gradients in the X and Y directions, respectively. R_{en}^X is the gradient of the dark light enhanced image in the X direction, and R_{gt}^X is the gradient of normal image in the X direction. \downarrow_D is the loss of the image local contrast difference.

3.3.2. Image Structural Loss and Minimum Absolute Deviation Loss

To smooth the image, this article adds the total variation loss \downarrow_S :

$$\downarrow_S = \frac{1}{W_{(i,j)}H_{(i,j)}} \sum_{i=1}^{W_{(i,j)}} \sum_{j=1}^{H_{(i,j)}} \{[(\sum_{R \in \sigma} |R_{(i,j)}^X|)/3]^2 - [(\sum_{R \in \sigma} |R_{(i,j)}^Y|)/3]^2\}, \sigma = \{r, g, b\}, \tag{14}$$

$\sum_{R \in \sigma} |R_{(i,j)}^X|$ represent the sum of the gradients of the red, green and blue channels in the X direction. $\sum_{R \in \sigma} |R_{(i,j)}^Y|$ represent the sum of the gradients of the red, green and blue channels in the Y direction. \downarrow_S is the image structural loss.

Absolute deviation loss is a commonly used loss function for calculating the loss before and after image enhancement:

$$\downarrow_1 = \sum_{i=1}^n (Y_i - X_i)^2. \tag{15}$$

where X_i is the original value and Y_i is the target value.

3.3.3. Image Integrity Loss

In a dark light enhanced network, two learned prior models are used as constraint functions, and a weighted sum is performed together with the defined local contrast difference measurement and structural loss of the image to obtain the total image loss, as shown in Formula (16):

$$\downarrow_T = \eta * \downarrow_D^{prior(LBP)} + \vartheta * \downarrow_D^{prior(MCMC)} + \mu * \downarrow_1 + \lambda * \downarrow_S \tag{16}$$

$\eta, \vartheta, \mu, \lambda$ represent weighting factors for the LBP prior loss $\downarrow_D^{prior(LBP)}$, MCMC prior loss $\downarrow_D^{prior(MCMC)}$, minimum absolute deviation loss \downarrow_1 and image structural loss \downarrow_S , respectively. \downarrow_T is the sum of losses in the process of enhancing low-light images.

4. Experiment

4.1. Implementation Details and Datasets

The publicly available datasets used for testing this algorithm included LOL, LIME, and DarkFace (the last two are both reference-free datasets). The evaluation indicators were divided into two parts: objective image indicators and subjective visual indicators. Among them, multiple self-encoding prior networks based on the LBP algorithm and MCMC algorithm were pre-trained. Then, the hyperparameter adjustment of multiple loss in the network was enhanced. The experimental hardware configuration was a PC with an Intel Xeon 4212R 2.40 GHz CPU, 128 GB RAM and an NVIDIA Quadro RTX 500 24GB GPU. The equipment comes from HP company in Beijing, China.

4.2. Ablation Experiment

This paper compares the performance of no prior network, LBP prior network, MCMC prior network and double prior network in the process of dark light image enhancement. LPIPS (learned perceptual image patch similarity), PSNR (peak signal to noise ratio) and SSIM (structural similarity index) are used as evaluation indexes to measure the difference of color, structure and detail. The higher the PSNR and SSIM values, the smaller the LPIPS value, and the better the image quality. The quantitative indicators on the LOL dataset are shown in Table 1.

Table 1. Image indexes of enhanced networks combined with different prior information.

Method	No Prior	MCMC Prior	LBP Prior	Double Priors
LPIPS	0.1877	0.1535	0.1428	0.1302
SSIM	0.7766	0.7998	0.7764	0.8122
PSNR(dB)	15.95	17.20	17.56	20.42

From Table 1, we can see that the enhancement effect of multiple prior collaborative constraints on dark light images is very obvious. Under the constraint of a single piece of prior information, the LPLPS value of the image-enhancement network decreases significantly, the SSIM is closer to 1 and the PSNR value is improved. The objective indicators are close to the better image, and the enhancement effect is more obvious under the collaborative constraint of double prior information. The enhancement effect of dark light enhancement networks with different priors on the LOL dataset is shown in Figure 3. Based

on the analysis of subjective visual perception variables such as detail clarity, contrast, color naturalness, brightness, etc., the results of the non-prior enhancement network have serious color distortion and large areas of artifacts. The enhancement network combined with the MCMC prior or LBP prior has improved in color and detail compared to the non-prior enhancement, but the results are still unsatisfactory. The image quality based on the MCMC prior are blurred in Figure 3h, and the image quality based on the LBP prior is too dark in the details of Figure 3d. Under a double prior co-constraint, the enhanced image is sharper in detail, with artifacts and noise interference suppressed, more natural colors, and is closer to the true value, as shown in Figure 3j.

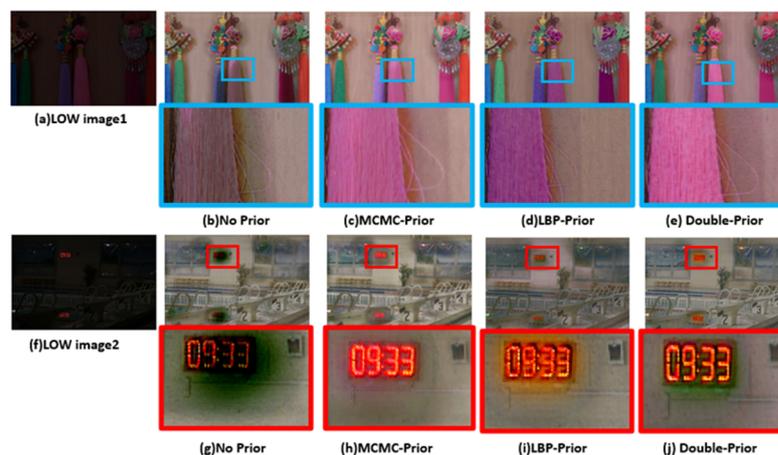


Figure 3. Detailed comparison results of dark light enhancement networks with different prior information on dark-light images of the LOL dataset. Among them, images (b–e) represent the different enhancement effects of image (a), while (g–j) represent the different enhancement effects of image (f). The four enhancement methods are in order: no prior enhancement, MCMC prior enhancement, LBP prior enhancement, and dual prior collaborative enhancement.

Information entropy can quantitatively enhance a network’s ability to compress data and learn potential missing information. A higher information entropy indicates that the random variable has more possibilities or more potential states, which represents the diversity of the optimized image. The 15 images of the LOL test set were enhanced with different prior information, and the information entropy of each image was calculated as shown in Figure 4. The image information entropy of the enhancement network without prior information is generally small. Most of the image information entropy in the enhancement results combined with the LBP prior or MCMC prior is better than that of the network without prior information, while the double-prior collaborative constraint network for dark light images performs well in the calculation of image information entropy.

4.3. Referenced Quality Assessment

This article uses the public dataset LOL to quantitatively evaluate various methods. These images have corresponding dark and true values. Therefore, the LOL dataset has more noise, the enhancement process is more complex, and the comparison of enhanced effects is more obvious. This article comprehensively evaluates the enhancement effect of images through PSNR, SSIM, NIQE (natural image quality evaluator) and LPIPS values. The performance of various dim-light enhancement algorithms on the LOL dataset is shown in Table 2. The test results of our algorithm on PSNR and SSIM are significantly better than those of other methods, and it also has a good ranking in the evaluation of LPIPS and NIQE indicators. Therefore, the enhancement approach proposed in this article is feasible and effective, and the comprehensive evaluation of objective indicators is superior to most existing methods.

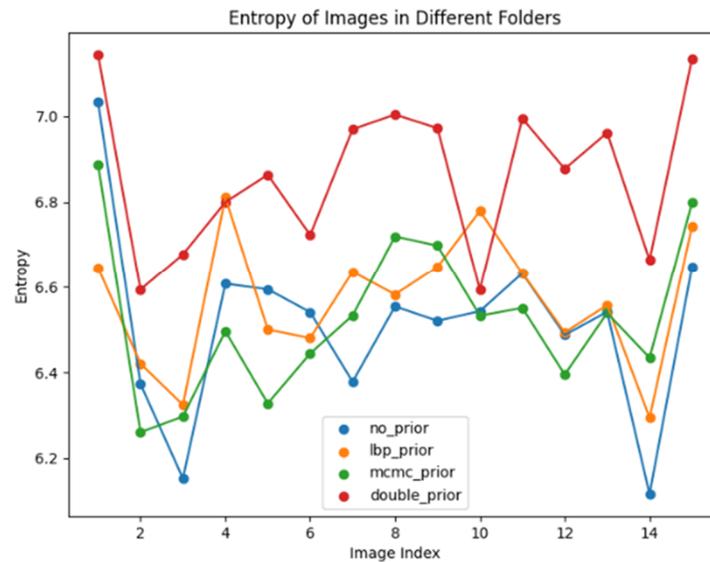


Figure 4. Comparison of information entropy of dark light enhancement networks with different prior information on 15 images of the LOL test set.

Table 2. Test results of various algorithms on the unified LOL dataset.

Method	PSNR (dB)	SSIM	LPIPS (alex)	LPIPS (vgg)	NIQE
DALE	17.39	0.750	0.0832	0.1243	15.054
DRBN	16.42	0.751	0.1197	0.2215	12.845
DSLRL	14.79	0.607	0.0861	0.1768	9.919
EnlightenGAN	17.50	0.666	0.1300	0.1743	10.001
RUAS	15.32	0.613	0.1440	0.2310	10.889
SGM	17.23	0.763	0.2820	0.3452	13.209
ZeroDCE	14.12	0.583	0.1362	0.1776	12.152
ZeroDCE++	14.37	0.589	0.1313	0.1689	11.876
KinD	16.44	0.789	0.1413	0.1695	9.658
KinD++	16.58	0.766	0.1590	0.1807	10.685
Ours	20.42	0.8122	0.1302	0.1665	10.922

The algorithm in this article has excellent performance in image enhancement quality, but in practical applications, processing time is also an effective reference for measuring the quality of the algorithm. In the collaborative constraint network based on multiple self-coding priors proposed in this article, although the model is relatively complex, all multiple priors are pre-trained models, which can be used as loadings to directly constrain and guide the enhancement process in the enhancement network, thus greatly saving training time. Because the training time is limited by the dataset and the number of cycles, it is not easy to effectively compare with other algorithms. This article is based on the models trained by various algorithms, and tests the enhancement speed on the LOL dataset’s test set (15 dim light images of 600×400). The test results are shown in Figure 5. The algorithm in this article has a processing time that exceeds most algorithms due to the limitation of model complexity. However, compared with other algorithms, the processing time difference of a single image is less than 0.2 s, which basically meets the real-time enhancement requirements and further improves the enhancement effect, fully proving the feasibility and superiority of the algorithm proposed in this paper.

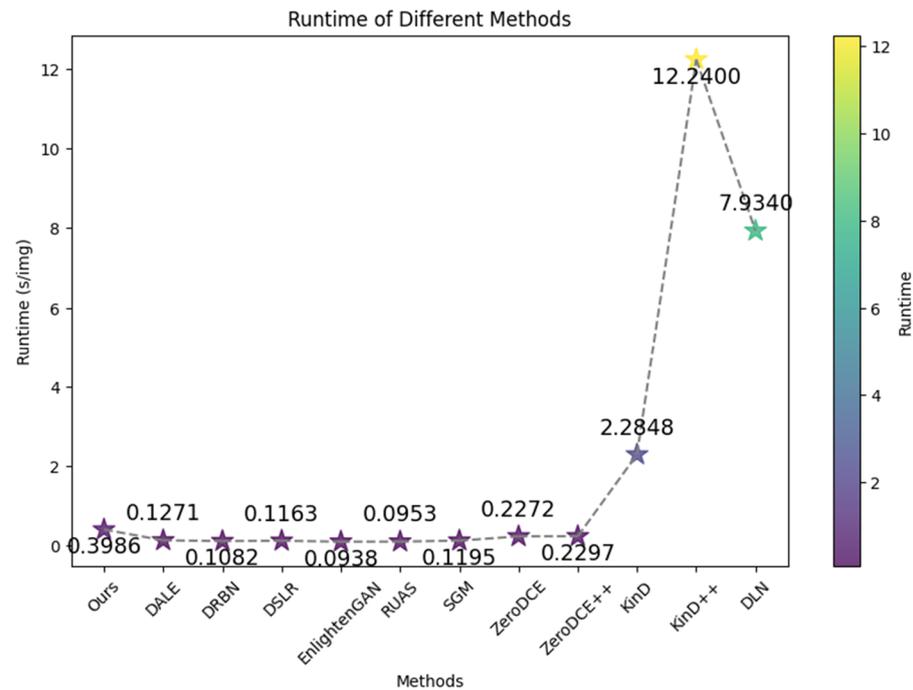


Figure 5. Processing speed map of multiple dim light enhancement algorithm models on a single 600×400 image.

The visual experience is shown in Figure 6. The objective evaluation indicators and intuitive visual perception are not completely unified. On the LOL dataset, the best ranked KinD algorithm in NIQE showed overall darkening, and changed the attributes of the image itself. For example, the shadow behind the rabbit doll image in Figure 6 was weakened, and the enhancement results of the DALE algorithm, which ranked first in the LPIPS metrics, were not satisfactory in terms of color. Due to the interference of noise and inconsistent exposure levels, most algorithms encounter problems such as color distortion, underexposure and blurred details during enhancement. However, the CCMP method proposed in this article does not exhibit distortion, overexposure or darkening in color brightness, and maintains good details and textures.

4.4. Quality Assessment without Reference

With the purpose is to verify the robustness of the method proposed, an enhancement test was conducted on a reference free public dataset. The visual performance on the reference dataset LIME is shown in Figure 7. It can be seen that most algorithms have poor robustness for images with different color tones and brightness levels. The images enhanced by the DALE, SGM and ZeroDCE algorithms generally exhibit underexposure, while the DLSR algorithm and RUAS algorithm exhibit bright blocks and derived regions. When enhancing the alien image, due to the high dynamic range characteristics and high contrast between light and dark, most algorithms experienced color imbalance during the enhancement process. Our algorithm (CCMP) has shown a great enhancement effect in both the overall color naturalness and details, while maintaining the original image properties in images with enhanced brightness. It can efficiently export high-quality images while suppressing noise and preserving details. Therefore, it can be proven that the more robust implicit priors learned through multiple self-coding modules have unique advantages in detail restoration, which testifies to the superiority of our collaborative constraint method.

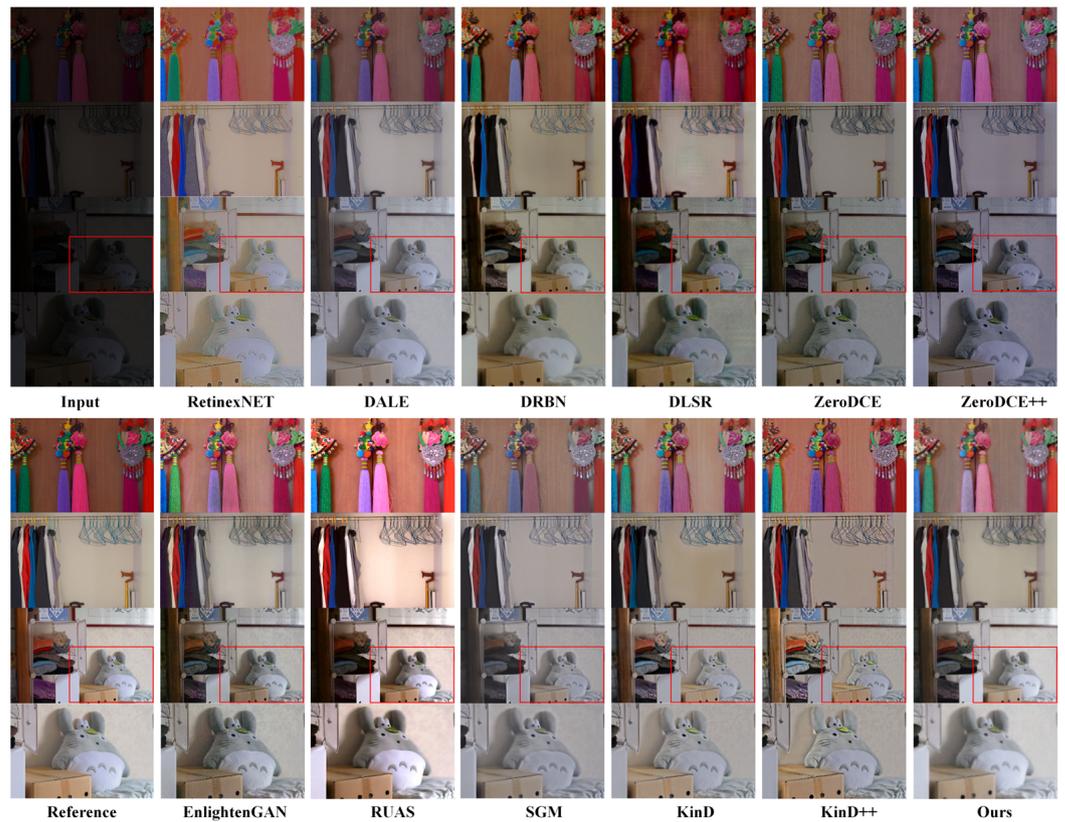


Figure 6. Comparative experiments of various dim light enhancement algorithms on the LOL test dataset.



Figure 7. Enhancement performance of various enhancement algorithms on the LIME dataset.

The CCMP algorithm was also compared on another publicly available dataset, the DarkFace dataset. As shown in Figure 8, on the DarkFace dataset, the enhanced images of the DALE, DLSR and ZeroDCE algorithms are all very dark, while the exported results of the DLSR, EnlightenGAN and RUAS algorithms show color deviation. After enhancement, our algorithm (CCMP) maintains the original color structure and has good clarity and details. Accordingly, the algorithm (CCMP) has certain advantages in enhancing performance in multi-scene dark environments.

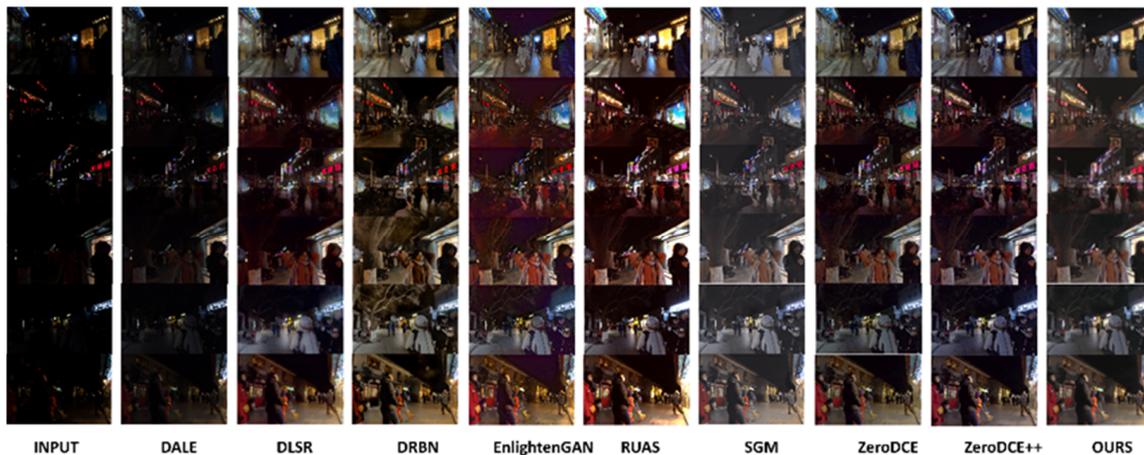


Figure 8. Enhancement results of dark light enhancement algorithms on the DarkFace dataset.

4.5. Experimental Discussion

Through the above tests, the superiority of the CCMP algorithm can be fully demonstrated through subjective visual perception and objective evaluation indicators. First, compared with the supervised network model representative algorithm DRBN (recursive band learning), the CCMP algorithm has a great enhancement effect, and the algorithm weakens the impact of the training process on the paired dataset. When trained on the same dataset and applied to different conditions of dark light images, the CCMP algorithm shows higher adaptability. This is because its unique multiple priors play a guiding role in the processing. Second, when comparing our algorithm with the classic unsupervised real-time enhancement algorithm represented by the Zero-DCE algorithm, the Zero-DCE algorithm is different from traditional deep learning algorithms that use large-scale network parameters for feature extraction. Instead, it uses several custom loss functions and neural networks to fit the brightness mapping curve, which is closer to mathematical calculations. However, due to sacrificing model complexity for processing speed, the Zero-DCE algorithm performed poorly on various datasets in Figures 5–7. However, although the CCMP algorithm increases the model complexity, it performs better in image processing with a wider range of effects. The processing time for a single image is only less than 0.2 s, and as the hardware level continues to improve, the time difference will further narrow. In addition, as for the EnlightenGAN algorithm, which is a deep-network model, although it does not rely on paired training data, EnlightenGAN uses generative adversarial networks to establish non paired mappings. Compared with the multiple prior information of the CCMP algorithm as guidance constraints, EnlightenGAN has greater randomness and lacks stability. Moreover, the multiple fidelity terms of CCMP also play a good balancing role, and the processing effect usually does not show significant deviations. In Figure 6, it can be seen that there is a significant deviation in EnlightenGAN, while the CCMP algorithm tends towards ideal truth values.

In summary, the CCMP algorithm demonstrates good adaptability to different environments due to its unique multiple prior information guidance. Additionally, multiple image-difference losses are designed as fidelity terms to ensure that the color, contrast and detail texture of the image do not deviate significantly from the ideal truth value, thus improving the stability of the algorithm. Additionally, directly loading prior information through pre-training saves processing time and does not slow down the processing speed due to the complexity of the model. Therefore, it confirms that the CCMP algorithm is innovative and progressive.

5. Conclusions

Aiming at the complex and changeable dark light environment, this article proposes a new approach for improving and optimizing images, which uses multiple specific autoencoder networks to mine implicit prior information from the original data, and uses the learned prior information in the form of collaborative constraints to guide the dark light image in the enhancement network. This includes features such as rich detail texture.

Experimental results on various low-light datasets evidence that our method outperforms many mainstream methods according to both subjective and objective indicators. Additionally, the structure of multiple prior information collaborative constraints can lead to better image quality improvement. Our future work will explore more effective prior learning methods under the premise of controlling the size of the model and obtaining additional feature information of the image, and try to add more enhancement constraint functions that meet the needs of the human eye to further optimize the effect of image enhancement.

Author Contributions: Conceptualization, J.H.; Methodology, H.W.; Formal analysis and investigation, Q.L.; Writing—original draft preparation, L.G.; Writing—review and editing, W.C. and J.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the West Light Foundation of the Chinese Academy of Sciences, grant No. XAB2021YN15.

Institutional Review Board Statement: This study does not require ethical approval.

Informed Consent Statement: This method is used for research that does not involve humans.

Data Availability Statement: The data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, W.; Wang, S.; Zhao, Y.; Tong, J.; Yang, T.; Li, D. Real-Time Obstacle Detection Method in the Driving Process of Driverless Rail Locomotives Based on DeblurGANv2 and Improved YOLOv4. *Appl. Sci.* **2023**, *13*, 3861. [[CrossRef](#)]
2. Zheng, X.; Zheng, S.; Kong, Y.; Chen, J. Recent Advances in Surface Defect Inspection of Industrial Products Using Deep Learning Techniques. *Int. J. Adv. Manuf. Technol.* **2021**, *113*, 35–58. [[CrossRef](#)]
3. Sreenu, G.; Saleem Durai, M.A. Intelligent Video Surveillance: A Review through Deep Learning Techniques for Crowd Analysis. *J. Big Data* **2019**, *6*, 48. [[CrossRef](#)]
4. Yuan, Q.; Shen, H.; Li, T.; Li, Z.; Li, S.; Jiang, Y.; Xu, H.; Tan, W.; Yang, Q.; Wang, J.; et al. Deep Learning in Environmental Remote Sensing: Achievements and Challenges. *Remote Sens. Environ.* **2020**, *241*, 111716. [[CrossRef](#)]
5. Feng, X.; Li, J.; Hua, Z.; Zhang, F. Low-Light Image Enhancement Based on Multi-Illumination Estimation. *Appl. Intell.* **2021**, *51*, 5111–5131. [[CrossRef](#)]
6. Tan, J.; Zhang, T.; Zhao, L.; Huang, D.; Zhang, Z. Low-Light Image Enhancement with Geometrical Sparse Representation. *Appl. Intell.* **2023**, *53*, 11019–11033. [[CrossRef](#)]
7. Wu, Y.; Song, W.; Zheng, J.; Liu, F. N2PN: Non-Reference Two-Pathway Network for Low-Light Image Enhancement. *Appl. Intell.* **2022**, *52*, 3559–3576. [[CrossRef](#)]
8. Land, E.H.; McCann, J.J. Lightness and Retinex Theory. *J. Opt. Soc. Am.* **1971**, *61*, 1. [[CrossRef](#)]
9. Jobson, D.J.; Rahman, Z.; Woodell, G.A. Properties and Performance of a Center/Surround Retinex. *IEEE Trans. Image Process.* **1997**, *6*, 451–462. [[CrossRef](#)]
10. Jobson, D.J.; Rahman, Z.; Woodell, G.A. A Multiscale Retinex for Bridging the Gap between Color Images and the Human Observation of Scenes. *IEEE Trans. Image Process.* **1997**, *6*, 965–976. [[CrossRef](#)]
11. Lee, C.-H.; Shih, J.-L.; Lien, C.-C.; Han, C.-C. Adaptive Multiscale Retinex for Image Contrast Enhancement. In Proceedings of the 2013 International Conference on Signal-Image Technology & Internet-Based Systems, Kyoto, Japan, 2–5 December 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 43–50.
12. Petro, A.B.; Sbert, C.; Morel, J.-M. Multiscale Retinex. *Image Process. Line* **2014**, *4*, 71–88. [[CrossRef](#)]
13. Wang, J.; He, N.; Lu, K. A New Single Image Dehazing Method with MSRRCR Algorithm. In Proceedings of the 7th International Conference on Internet Multimedia Computing and Service, Zhangjiajie, China, 19–21 August 2015; ACM: New York, NY, USA, 2015; pp. 1–4.

14. Wang, J.; Lu, K.; Xue, J.; He, N.; Shao, L. Single Image Dehazing Based on the Physical Model and MSRCR Algorithm. *IEEE Trans. Circuits Syst. Video Technol.* **2018**, *28*, 2190–2199. [[CrossRef](#)]
15. Lu, K.; Zhang, L. TBEFN: A Two-Branch Exposure-Fusion Network for Low-Light Image Enhancement. *IEEE Trans. Multimed.* **2021**, *23*, 4093–4105. [[CrossRef](#)]
16. Tao, L.; Zhu, C.; Xiang, G.; Li, Y.; Jia, H.; Xie, X. LLCNN: A Convolutional Neural Network for Low-Light Image Enhancement. In Proceedings of the 2017 IEEE Visual Communications and Image Processing (VCIP), St. Petersburg, FL, USA, 10–13 December 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–4.
17. Wang, W.; Wei, C.; Yang, W.; Liu, J. GLADNet: Low-Light Enhancement Network with Global Awareness. In Proceedings of the 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, China, 15–19 May 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 751–755.
18. Hai, J.; Xuan, Z.; Han, S.; Yang, R.; Hao, Y.; Zou, F.; Lin, F. R2RNet: Low-Light Image Enhancement via Real-Low to Real-Normal Network. *J. Vis. Commun. Image Represent.* **2023**, *90*, 103712. [[CrossRef](#)]
19. Li, J.; Feng, X.; Hua, Z. Low-Light Image Enhancement via Progressive-Recursive Network. *IEEE Trans. Circuits Syst. Video Technol.* **2021**, *31*, 4227–4240. [[CrossRef](#)]
20. Wei, C.; Wang, W.; Yang, W.; Liu, J. Deep Retinex Decomposition for Low-Light Enhancement. *arXiv* **2018**, arXiv:1808.04560.
21. Wu, W.; Weng, J.; Zhang, P.; Wang, X.; Yang, W.; Jiang, J. URetinex-Net: Retinex-Based Deep Unfolding Network for Low-Light Image Enhancement. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 5891–5900.
22. Yang, W.; Wang, W.; Huang, H.; Wang, S.; Liu, J. Sparse Gradient Regularized Deep Retinex Network for Robust Low-Light Image Enhancement. *IEEE Trans. Image Process.* **2021**, *30*, 2072–2086. [[CrossRef](#)] [[PubMed](#)]
23. Shen, L.; Yue, Z.; Feng, F.; Chen, Q.; Liu, S.; Ma, J. MSR-Net: Low-Light Image Enhancement Using Deep Convolutional Network. *arXiv* **2017**, arXiv:1711.02488.
24. Liu, R.; Ma, L.; Zhang, J.; Fan, X.; Luo, Z. Retinex-Inspired Unrolling with Cooperative Prior Architecture Search for Low-Light Image Enhancement. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 10556–10565.
25. Ren, W.; Liu, S.; Ma, L.; Xu, Q.; Xu, X.; Cao, X.; Du, J.; Yang, M.-H. Low-Light Image Enhancement via a Deep Hybrid Network. *IEEE Trans. Image Process.* **2019**, *28*, 4364–4375. [[CrossRef](#)]
26. Fan, M.; Wang, W.; Yang, W.; Liu, J. Integrating Semantic Segmentation and Retinex Model for Low-Light Image Enhancement. In Proceedings of the 28th ACM International Conference on Multimedia, Seattle, WA, USA, 12–16 October 2020; ACM: New York, NY, USA, 2020; pp. 2317–2325.
27. Liu, L.; An, J.; Liu, J.; Yuan, S.; Chen, X.; Zhou, W.; Li, H.; Wang, Y.F.; Tian, Q. Low-Light Video Enhancement with Synthetic Event Guidance. In Proceedings of the AAAI Conference on Artificial Intelligence, Washington, DC, USA, 7–14 February 2023; Volume 37, pp. 1692–1700. [[CrossRef](#)]
28. Wang, Y.; Cao, Y.; Zha, Z.-J.; Zhang, J.; Xiong, Z.; Zhang, W.; Wu, F. Progressive Retinex: Mutually Reinforced Illumination-Noise Perception Network for Low Light Image Enhancement. In Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 21–25 October 2019.
29. He, K.; Chen, X.; Xie, S.; Li, Y.; Dollar, P.; Girshick, R. Masked Autoencoders Are Scalable Vision Learners. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 15979–15988.
30. Chen, X.; Ding, M.; Wang, X.; Xin, Y.; Mo, S.; Wang, Y.; Han, S.; Luo, P.; Zeng, G.; Wang, J. Context Autoencoder for Self-Supervised Representation Learning. *Int. J. Comput. Vis.* **2023**, *132*, 208–223. [[CrossRef](#)]
31. Dong, X.; Bao, J.; Zhang, T.; Chen, D.; Zhang, W.; Yuan, L.; Chen, D.; Wen, F.; Yu, N.; Guo, B. PeCo: Perceptual Codebook for BERT Pre-Training of Vision Transformers. In Proceedings of the AAAI Conference on Artificial Intelligence, Washington, DC, USA, 7–14 February 2023; Volume 37, pp. 552–560. [[CrossRef](#)]
32. Wei, C.; Fan, H.; Xie, S.; Wu, C.-Y.; Yuille, A.; Feichtenhofer, C. Masked Feature Prediction for Self-Supervised Visual Pre-Training. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 14648–14658.
33. Assran, M.; Caron, M.; Misra, I.; Bojanowski, P.; Bordes, F.; Vincent, P.; Joulin, A.; Rabbat, M.; Ballas, N. Masked Siamese Networks for Label-Efficient Learning. In *European Conference on Computer Vision*; Springer Nature: Cham, Switzerland, 2022.
34. Prat, A.; Sautory, T.; Navarro-Martinez, S. A Priori Sub-Grid Modelling Using Artificial Neural Networks. *Int. J. Comput. Fluid Dyn.* **2020**, *34*, 397–417. [[CrossRef](#)]
35. E, W.; Ma, C.; Wu, L. A Priori Estimates of the Population Risk for Two-Layer Neural Networks. *Commun. Math. Sci.* **2019**, *17*, 1407–1425. [[CrossRef](#)]
36. Huang, W.; Huang, Y.; Wu, Z.; Yin, J.; Chen, Q. A Multi-Kernel Mode Using a Local Binary Pattern and Random Patch Convolution for Hyperspectral Image Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 4607–4620. [[CrossRef](#)]
37. Gabrié, M.; Rotskoff, G.M.; Vanden-Eijnden, E. Adaptive Monte Carlo Augmented with Normalizing Flows. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2109420119. [[CrossRef](#)]

-
38. Karamanis, M.; Beutler, F.; Peacock, J.A.; Nabergoj, D.; Seljak, U. Accelerating Astronomical and Cosmological Inference with Preconditioned Monte Carlo. *Mon. Not. R. Astron. Soc.* **2022**, *516*, 1644–1653. [[CrossRef](#)]
 39. Nemeth, C.; Fearnhead, P. Stochastic Gradient Markov Chain Monte Carlo. *J. Am. Stat. Assoc.* **2021**, *116*, 433–450. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.