



Changlin Wu<sup>1,2</sup>, Dandan Wang<sup>1,2,\*</sup>, Kaifeng Huang<sup>1,2</sup> and Long Wu<sup>1,2</sup>

- <sup>1</sup> School of Mechanical and Electrical Engineering, Huainan Normal University, Huainan 232001, China; wuchanglin@hnnu.edu.cn (C.W.); hkf@hnnu.edu.cn (K.H.); lwu@hnnu.edu.cn (L.W.)
- <sup>2</sup> Human-Computer Collaborative Robot Joint Laboratory of Anhui Province, Hefei 230002, China

Correspondence: lansejingling1988@126.com

Abstract: The dim lighting and excessive dust in underground mines often result in uneven illumination, blurriness, and loss of detail in surveillance images, which hinders subsequent intelligent image recognition. To address the limitations of the existing image enhancement algorithms in terms of generalization and accuracy, this paper proposes an unsupervised method for enhancing mine images in the hue-saturation-value (HSV) color space. Inspired by the HSV color space, the method first converts RGB images to the HSV space and integrates Retinex theory into the brightness (V channel). Additionally, a random perturbation technique is designed for the brightness. Within the same scene, a U-Net-based reflectance estimation network is constructed by enforcing consistency between the original reflectance and the perturbed reflectance, incorporating ResNeSt blocks and a multi-scale channel pixel attention module to improve accuracy. Finally, an enhanced image is obtained by recombining the original hue (H channel), brightness, and saturation (S channel), and converting back to the RGB space. Importantly, this image enhancement algorithm does not require any normally illuminated images during training. Extensive experiments demonstrated that the proposed method outperformed most existing unsupervised low-light image enhancement methods, qualitatively and quantitatively, achieving a competitive performance comparable to many supervised methods. Specifically, our method achieved the highest PSNR value of 22.18, indicating significant improvements compared to the other methods, and surpassing the second-best WCDM method by 10.3%. In terms of SSIM, our method also performed exceptionally well, achieving a value of 0.807, surpassing all other methods, and improving upon the second-place WCDM method by 19.5%. These results demonstrate that our proposed method significantly enhanced image quality and similarity, far exceeding the performance of the other algorithms.

Keywords: retinex; mine images; U-Net; HSV; ResNeSt

# 1. Introduction

Intelligent visual recognition technologies are widely applied in smart mining [1]. However, challenges such as insufficient lighting, uneven illumination, and dust in mine environments significantly affects the performance of image analysis techniques, including target tracking [2], object recognition [3], image semantic segmentation [4], and object detection [5]. To obtain high-quality mine images, it is crucial to effectively enhance low-light images [6] by improving brightness, illumination uniformity, color balance, and detail information. Therefore, researching methods suitable for mine image enhancement is of great importance [7].

Traditional image enhancement algorithms for addressing the low-light environments in coal mines include histogram equalization [8,9], dark channel prior enhancement [10], the Retinex algorithm [11], and image enhancement based on the hue–saturation–value (HSV) color space [12,13]. Among these, the Retinex algorithm is widely used and encompasses various techniques such as single scale Retinex (SSR) [14], multiScale Retinex (MSR) [15],



Citation: Wu, C.; Wang, D.; Huang, K.; Wu, L. Enhancement of Mine Images through Reflectance Estimation of *V* Channel Using Retinex Theory. *Processes* **2024**, *12*, 1067. https:// doi.org/10.3390/pr12061067

Academic Editor: Amin Beiranvand Pour

Received: 28 April 2024 Revised: 19 May 2024 Accepted: 22 May 2024 Published: 23 May 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/). and color restoration Retinex (MSRCR) [16]. The fundamental principle of Retinex theory is based on the assumption that the color of an object remains constant regardless of variations in illumination, exhibiting color constancy. As a result, the Retinex algorithm decomposes the input image into illumination and reflectance components. The reflectance component represents the inherent properties of objects, remaining unaffected by the illumination conditions, while the illumination component captures the differences between low and normal lighting.

With the continuous advancement of technology, deep learning has demonstrated significant potential for applications in various fields. It is increasingly being employed in the domain of low-light image enhancement [17]. Based on the differences in methodology, these approaches can be categorized into supervised and unsupervised methods.

Supervised methods require collecting paired images in advance to establish the mapping relationship between normal and low-light images [18–34]. Zhang et al. [18] proposed the KinD++ algorithm, which makes the processed images more realistic. Jiang et al. [20] introduced a novel degradation to refinement generative network (DRGN) that focuses more on the detailed information of images, resulting in improved visibility, while maintaining naturalness. Xu et al. [21] proposed a hierarchical feature mining network called HFMNet, which extracts illumination and edge features from different network layers. They also developed a feature mining attention (FMA) module and utilized hierarchical supervision loss to fully exploit key features, thereby enhancing image visibility. Wu et al. [25] presented URetinex Net, a deep unrolling network based on Retinex theory, which combines an implicit prior regularization model with Retinex theory to better suppress noise and preserve details. Lu et al. [27] proposed a frequency-divided multiscale learning network that decomposes low-light images into high-frequency and low-frequency components. They employed multiscale feature extraction and attention mechanisms to enhance the images. Chen et al. [30] combined convolutional and self-attention mechanisms to enhance low-light images, focusing on both local and global features during feature extraction to eliminate noise and color biases. Yang et al. [31] introduced an enhancement network based on multi-stream information supplementation, which restores the rich structure of images through information supplementation, while removing noise and color biases. Xu et al. [32] proposed a method called SNR-aware low-light image enhancement (SNRNet), which utilizes a signal-to-noise ratio-aware transformer and a convolutional model to dynamically enhance pixels with spatially varying operations. Jiang et al. [33] introduced a waveletbased conditional diffusion model (WCDM) that leverages the generative capability of the diffusion model to generate perceptually faithful results. Li et al. [34] presented a novel solution called UHDFour, which incorporates Fourier transform into a cascaded network.

Supervised methods rely on paired datasets to construct end-to-end networks that learn the mapping relationship between low-light and normal-light images. These methods heavily rely on feature extraction, and attention mechanisms enable the networks to focus on more informative features. Designing an appropriate network allows for better expression of the intrinsic information of images, resulting in high-quality image generation. However, there are still limitations to this approach, which can be summarized as follows:

- (1) Supervised methods depend on paired datasets, which are challenging and timeconsuming to acquire.
- (2) The generalization capability of the models is insufficient. Since they rely on paired datasets for training, they may achieve impressive results on specific datasets but perform poorly in other scenarios.

Unsupervised methods do not require paired datasets [35–47], which greatly enhances the research efficiency. Guo et al. [36] proposed a novel unsupervised deep curve estimation (Zero DCE) approach, utilizing a lightweight deep network called DCE-Net to adjust the dynamic range of given images, without the need for preprepared paired datasets. However, this method does not consider the influence of noise and suffers from a larger network size, resulting in increased computational times. Jiang et al. [38] introduced a fast and effective enhancement strategy (EnlightenGAN) by combining an attention

mechanism with a U-Net network. Fu et al. [39] presented a new unsupervised low-light image enhancement network, LE-GAN, based on the generative adversarial network (GAN) framework and trained using non-paired low-light images. This network incorporates a perceptual attention module for enhanced feature extraction, noise removal, and color bias elimination. It further introduces novel loss functions to address overexposure issues. Qiao et al. [42] proposed a generative adversarial network approach based on reverse attention mechanisms, utilizing a deep aggregation pyramid pooling module to remove noise and a reverse attention module to eliminate color aberrations and enhance image quality. Jiang et al. [44] introduced an unsupervised decomposition and correction network, incorporating a noise removal network to suppress noise. Inspired by CycleGAN [45], Bhattacharya et al. [46] proposed an image enhancement model (D2BGAN) that effectively removes artifacts by combining geometric and illumination consistency, contextual loss, and multi-scale color, texture, and edge discriminators. Ma et al. [47] introduced a novel self-calibrated illumination (SCI) learning framework designed to enhance the brightness of low-light images in real-world scenarios with speed, flexibility, and robustness.

Unsupervised methods, by not requiring paired datasets, offer a more focused exploration of the relationship between low-light and normal-light images. While they eliminate the need for paired data, the resulting image quality often falls short of expectations and exhibits certain visual artifacts. As demonstrated in Figure 1, unsupervised methods generally yield lower image enhancement quality compared to the majority of supervised methods.



Figure 1. Using the LOL dataset, a comparison of the state-of-the-art supervised and unsupervised image enhancement methods in terms of peak signal-to-noise ratio (PSNR) is presented. It should be noted that a

enhancement methods in terms of peak signal-to-noise ratio (PSNR) is presented. It should be noted that a higher PSNR indicates better enhancement results. It is evident that the majority of supervised methods consistently outperform unsupervised methods in terms of image enhancement effects.

In response to the issues presented by the aforementioned algorithms, this paper proposes an unsupervised image enhancement method for mining images based on the HSV color space. Firstly, we convert the low-light RGB images captured underground to the HSV space, and then combine them with Retinex theory to design a U-Net-based reflectance estimation network. Our ultimate objective is to improve the image quality generated by unsupervised methods. Extensive experiments conducted on our self-constructed dataset demonstrated that our method surpassed the state-of-the-art (SOTA) methods in terms of distortion and perceptual metrics. The contributions of our work can be summarized as follows:

- (1) Combination of Retinex Theory and HSV Color Space for Reflectance Estimation: In our study, we propose an innovative approach to enhancing mining images by combining Retinex theory with reflectance estimation in the *V* channel of the HSV color space. By utilizing a deep reflectance estimation network, we accurately estimate the reflectance information in the images, leading to a significant improvement in visibility and image quality.
- (2) Design of an Enhanced U-Net Model: We introduce an improved U-Net model that integrates the U-Net network, ResNeSt blocks, and a multi-scale channel-wise pixel attention mechanism. Through operations such as convolution, encoding downsampling, feature duplication, and upsampling, we achieve enhancement in the *V* channel. This model effectively extracts image features and applies attention modules in the decoder to retain important information, resulting in outstanding enhancement performance.
- (3) Introduction of Four Sets of Loss Functions for Model Optimization: To address the challenge of enhancing low-light images in underground coal mines, we design four sets of loss functions, including an exposure loss function, reflectance consistency loss function, spatial structure loss function, and illumination smoothness loss function. By incorporating these loss functions, we comprehensively consider the exposure level, reflectance consistency, spatial structure details, and illumination smoothness of the images, leading to holistic model optimization and improved performance in low-light image enhancement.
- (4) Introduction of a Random Brightness Perturbation Method: Our study introduces a novel random brightness perturbation method to enhance the robustness and generalization ability of the model. By introducing randomness during training, we can simulate the variations in reflectance under different lighting conditions in real-world scenarios. This enables our model to adapt to various situations and demonstrate better performance in practical applications.

# 2. Image Enhancement Network Model Based on HSV Space Transformation

## 2.1. Overall Architecture of the Model

This paper presents a mine image enhancement model based on the HSV color space, as illustrated in Figure 2. Initially, the mine images in the RGB space are transformed into the HSV color space, and the hue (*H* channel), saturation (*S* channel), and value (*V* channel) components are extracted. While keeping the *H* and *S* components unchanged, we integrate a Retinex theory model with the luminance component *V* and employ an improved U-Net network to enhance the luminance. Additionally, by introducing a random perturbation method for the luminance, a perturbation component *V*' is obtained. To enhance the model's generalization capability, in the same scene, we increase the model's generalization by enforcing consistency between the original reflectance and the new randomly perturbed reflectance. The random perturbation method used for the luminance enables the network model to achieve effective image enhancement, even in the absence of normal images. The enhanced luminance components  $V_c$  and  $V'_c$ . Finally, the enhanced HSV results are transformed back into the RGB space, resulting in the output of an enhanced image.



**Figure 2.** Introduction to Our Overall Model Architecture: (a) HSV Color Transformation: Starting from RGB images as input, this component generates HSV features as output. (b) Enhanced U-Net: This module integrates the U-Net network, ResNeSt block, and multi-scale channel pixel attention mechanism to enhance the *V* channel. (c) RGB Color Transformation: The HSV representation is subsequently converted back to the RGB space through this transformation.

# 2.2. The HSV Color Space

HSV is a widely used color system in everyday life, as it provides a more intuitive way for people to describe colors, including the hue (H), saturation (S), and value (V) components [48]. The hue represents the color tone using degrees, while the saturation refers to the intensity or depth of the color. The value component corresponds to the brightness of the color. The HSV space exhibits the advantage of component independence, enabling flexible adjustments of different attributes, without significant interference from

the other components. In particular, the brightness component (V channel) in the HSV space reflects the image's lightness. Enhancing the brightness enhances the contrast, details, and visibility of contours in the image. Therefore, converting an image to the HSV color space allows for a partitioning of the image into H and S components, which are closely related to color, while enabling specific adjustments to be made solely on the V channel to control image brightness. The conversion formula from the RGB color space to the HSV color space is as follows [49]:

$$H = \begin{cases} 0^{\circ} & \Delta = 0\\ 60^{\circ} \times \left(\frac{G' - B'}{\Delta} + 0\right) & , C \max = R'\\ 60^{\circ} \times \left(\frac{B' - R'}{\Delta} + 2\right) & , C \max = G'\\ 60^{\circ} \times \left(\frac{R' - G'}{\Delta} + 4\right) & , C \max = B' \end{cases}$$
(1)

$$S = \begin{cases} 0 & C \max = 0\\ \frac{\Delta}{C \max} & C \max \neq 0 \end{cases}$$
(2)

$$V = C \max$$
 (3)

Among them, R', G', and B' normalize the values of R, G, and B to  $0 \sim 1$ . C max is the maximum value in R', G', B', and  $\Delta$  is the difference between C max and C min.

Moreover, we conducted a straightforward experiment to demonstrate the advantages of the HSV color space. We randomly selected a pair of low-light/normal-light images from a dataset and transformed them into the HSV color space. By concatenating the *H* and *S* channels of the low-light image with the *V* channel of the corresponding normal-light image, we created a recombined image. Subsequently, we converted the recombined image back to the RGB color space. As depicted in Figure 3, the recombined image exhibited a color fidelity similar to that of the normal-light image. Hence, for low-light image enhancement based on the HSV color space, it suffices to enhance the brightness, as the color information is solely derived from the low-light image.



(a)Normal mine image (b)Low light mine image

(c) Reconstructed image

**Figure 3.** Visual quality results of the recomposed images generated from the H and S channels of low-light images and the corresponding *V* channel of the respective bright images.

# 2.3. Random Brightness Interference

To simulate the reflection differences caused by lighting variations in real-world scenarios, this paper proposes a brightness random perturbation method based on a power function. This method adjusts the image brightness by introducing a random exponent, thereby generating new images with the same reflectance characteristics as the original image but with different brightness levels. This random perturbation method offers the following advantages:

(1) Range Control: Both the input and output of the power function fall within the range [0, 1], avoiding information loss due to the brightness exceeding a reasonable range.

- (2) Monotonicity Preservation: The power function is monotonically increasing, ensuring consistent brightness relationships between the original and perturbed images, preventing unnatural brightness changes.
- (3) Perturbation Diversity: By introducing a random exponent, various degrees of brightness perturbations can be generated, increasing the diversity of training data and aiding in improving the robustness of the model.

$$V'(x) = V(x)^d \tag{4}$$

In this context, V(x) represents the pixel values of the original *V*-channel, V'(x) denotes the perturbed *V*-channel, and *d* represents a random value. In ablation experiments, we validated the effectiveness of this power function-based random brightness perturbation method and demonstrated its ability to accurately simulate the reflection differences caused by lighting variations in real-world scenarios. Furthermore, we observed that it had a positive impact on enhancing the performance of the image enhancement algorithms.

# 2.4. Retinex Theory

Most Retinex-based networks are initially trained with a decomposition network, to learn the decomposition of the illumination and reflectance components [50]. However, this process inevitably introduces reconstruction errors when generating the final image. In this study, we propose a network architecture that incorporates a perturbation, to introduce an interfering brightness, utilizing the same network structure with shared parameters. According to Retinex theory [51], an image can be decomposed into brightness and reflectance components, where the reflectance component should remain invariant under different lighting conditions. Hence, the aforementioned channels, denoted as V and its perturbed counterpart V' can be decomposed into the following two parts.

$$\begin{cases} V = R \times L \\ V' = R' \times L' \end{cases}$$
(5)

In the given equation, where R represents the reflectance component and L represents the illumination component, we consider L an intermediate variable for computing R. Hence, the above equation can be reformulated as follows:

$$\begin{cases} R = V \times \frac{1}{L} \\ R' = V' \times \frac{1}{L'} \end{cases}$$
(6)

By considering this approach, it becomes unnecessary to compute the difference between *V* and R \* L, thereby avoiding information loss during image reconstruction [52].

# 2.5. An Improved U-Net-Based Network for Reflectance Estimation

The U-Net network architecture possesses the characteristics of a deep convolutional neural network [53], enabling effective extraction of feature information from images. It adopts an encoder–decoder structure, utilizing multiple layers of convolution and pooling operations to capture both the fine details and contextual information in an image, thereby enhancing the effectiveness of image enhancement. However, the U-Net network may encounter challenges when dealing with complex image structures and textures. Due to its simple encoder–decoder structure, U-Net may struggle to accurately restore intricate details and textures in more complex scenes, leading to distortion or blurring in the enhanced results.

In this study, to meet our application requirements and ensure that the enhanced results better preserve image information, we propose a lightweight multi-scale residual attention U-Net model by integrating the advantages of the traditional U-Net network structure with ResNeSt blocks and multi-scale channel pixel attention modules. The architecture of this model is illustrated in Figure 2. We replace a portion of the  $3 \times 3$  convolution modules in the U-Net model with ResNeSt blocks (as depicted in Figure 4), thereby increas-

ing the network's width and its adaptability to multi-scale targets. To address the issue of the U-Net model's skip connections propagating irrelevant information or noise along with crucial detailed features to the decoding layers, we introduce attention mechanisms into the network model (as shown in Figure 5).



Figure 4. ResNeSt block. A detailed view of the split-attention unit is shown in Figure 5.



Figure 5. Split-attention within a cardinal group.

#### 2.5.1. ResNeSt Block

In Figure 2, at each level of the U-Net network, the feature tensors sequentially pass through a  $3 \times 3$  convolution block and a ResNeSt block (indicated by the red arrows in Figure 2). The ResNeSt [54] block retains the unique residual structure of ResNet, allowing the network to be deep, without suffering from the degradation problem associated with deep networks. Furthermore, to select the more crucial information from a plethora of features, the model incorporates an attention mechanism inspired by SENet [55] (squeeze and excitation networks) and SKNet [56] (selective kernel network) when designing the feature extraction units. The structure, as shown in Figure 4, divides the input into k cardinals, each denoted Cardinal i. Each cardinal's base array is split into r parts, and each part undergoes  $1 \times 1$  and  $3 \times 3$  convolution operations. After extracting the feature information, the split r-part module information is fed into the split attention module, achieving an output of base array information. The results of each group are then concatenated. Finally, the aggregated output feature map is linearly combined with the feature map of the residual module to obtain the final output.

After performing a 1 × 1 convolution operation on the input features, a grouping operation is implemented. The input is divided into *K* base arrays, and within each base array, the features are further split into *R* parts, resulting in a total of G = KR features. Within each sub-base array, the input image undergoes feature extraction operations, resulting in output features denoted  $U_i = F_i(x)$ ,  $i \in \{1, 2, ...KR\}$ . Subsequently, the features from the same base group are input into the split-attention module, as illustrated in Figure 4, where feature summation takes place, as expressed by Equation (7).

$$U^{k} = \sum_{j=R(k-1)+1}^{KR} U_{j}$$
(7)

In this equation [54],  $U^k \in N^{H \times W \times C/K}$ ,  $k \in \{1, 2, ..., K\}$  represents the sum of features in the *k*-th base group.  $U_j$  represents the feature information obtained after the  $1 \times 1$ and  $3 \times 3$  convolution operations on the *j*-th split block within the *k*-th unit group,  $j \in \{Rk - 1 + 1, R(k - 1) + 2, ..., Rk\}$ . Subsequently, it undergoes global average pooling across spatial dimensions, to capture the global context information embedded in the channels. The calculation expression for the *c*-th channel is given by Equation (8).

$$S_{\rm c}^{\rm k} = \frac{1}{H \times W} \sum_{\rm i}^{\rm H} \sum_{\rm j}^{\rm W} U_{\rm C}^{\rm K} {\rm i}, {\rm j}$$
(8)

Lastly, the slices of the *c*-th dimension channel are aggregated using weighted summation, denoted as  $V_c^k$ , as expressed in Equation (9).

$$V_{\rm c}^{\rm k} = \sum_{\rm i}^{\rm R} {\rm a}_{i}^{\rm k}(c) U_{{\rm R}(k-1)+i} \tag{9}$$

In this equation,  $a_i^k(c)$  represents the allocation weights, expressed as shown in Equation (10). Here,  $\xi_i^c$  represents the weights for each segmentation in the *C* channels.

$$\mathbf{a}_{i}^{k}(c) = \begin{cases} \frac{\exp(\xi_{i}^{c}(s^{k}))}{\sum\limits_{j=1}^{R} \exp(\xi_{i}^{c}(s^{k}))}, R > 1\\ \frac{1}{1 + \exp(\xi_{i}^{c}(s^{k}))}, R = 1 \end{cases}$$
(10)

Finally, all unit groups are connected along the channel dimension. After the split attention operation, a feature map V is obtained. Combining this with the ResNet network, the input features are skip-connected with the output features, resulting in the final output. This process enables the capture of more comprehensive local geometric

details and generates more complete high-level semantic information, facilitating deeper information mining.

## 2.5.2. Multi-Scale Channel Pixel Attention Module

In the conventional U-Net network, the decoding process in the decoder tends to lose important fine-grained details. To address this issue, skip connections are introduced in the U-Net architecture to establish connections between the encoder and decoder [53]. These skip connections enable the transfer of feature information extracted from the encoder to the decoder. However, such a structure can be susceptible to the negative impact of redundant information from the encoder on the feature fusion in the decoder. To overcome this problem, attention mechanisms can be incorporated to alleviate the issue and enhance the feature fusion process.

A multi-scale channel pixel attention module is designed, as illustrated in Figure 6. The module achieves multi-scale functionality by concurrently utilizing  $3 \times 3$ ,  $5 \times 5$ , and  $7 \times 7$  convolutions. In general, larger convolution kernels result in larger receptive fields, allowing the network to better capture global features in the image. Conversely, smaller convolution kernels have smaller receptive fields, enabling the network to extract finer local features.



Figure 6. Multi-scale channel pixel attention module.

After convolution, the input image generates multi-channel features, with each channel representing different feature information within the image. Moreover, the different feature types of information hold varying degrees of importance in enhancing low-light images. Therefore, channel attention is employed to capture the correlation among different feature channels. Specifically, the multi-channel features extracted through convolution are compressed into channel descriptors using adaptive average pooling, which encodes the global spatial information of the feature maps. Subsequently, by applying convolution– ReLU–convolution–sigmoid operations, feature weight matrices for different channels are obtained. Finally, the input features are multiplied by their respective weight matrices to obtain the output features. Following the channel attention module, a pixel attention module is added to help the network focus on informative image features, such as high-frequency regions and areas with severe blurriness. Specifically, the weighted features obtained from the channel attention module serve as input to the pixel attention module. Through convolution–ReLU–

convolution–sigmoid operations, weights for different pixel values are generated. These weights are then multiplied with the input features to obtain new features. To fully utilize the information from different scales, the output features of the three

branches  $(3 \times 3, 5 \times 5, \text{ and } 7 \times 7)$  are concatenated, and convolution is applied to fuse the three features. This fused feature is then passed to the next encoding and the corresponding decoding layers. Additionally, residual connections are incorporated by adding the input features to the output of the multi-scale channel pixel attention module. This helps the network learn more stable features.

### 2.6. Design of the Loss Function

To address the enhancement of low-light images in underground coal mines, four sets of loss functions are employed for model optimization. These include the exposure loss function, reflectance consistency loss function, spatial structure loss function, and illumination smoothness loss function.

(1) Exposure Loss: In order to control the brightness of the generated reflectance map, the difference between the average brightness of the reflectance and the given exposure value of a normal image is computed. Let *E* be set to 0.7, then the exposure loss  $L_1$  can be defined as follows:

$$L_1 = \|R - E\|_2^2 \tag{11}$$

where *R* represents the generated reflectance and *E* is a matrix of the same dimensions as *R*, with all values set to 0.7.

(2) Reflectance Consistency Loss: According to Retinex theory, the reflectance of *V* should be consistent with V'. Therefore, the reflectance consistency loss  $L_2$  can be defined as follows [36]:

$$L_2 = \|R - R'\|_2^2 \tag{12}$$

where R and R' are the reflectance maps generated from V and V', respectively, and  $|||_2$  represents the  $L_2$  norm.

(3) Spatial Structure Loss: To preserve the spatial structure of the input image in the reflectance map, a spatial structure loss is introduced. This loss measures the difference between the horizontal and vertical gradients of each pixel between the input image and the reflectance map. The spatial structure loss  $L_3$  can be defined as follows:

$$L_3 = \|\nabla \mathbf{R}_{\mathbf{m}} - \nabla \mathbf{V}_{\mathbf{m}}\|_2^2 \tag{13}$$

where  $\mathbf{R}_{\mathbf{m}}$  and  $\mathbf{V}_{\mathbf{m}}$  represent the generated reflectance and the average pooling result of the input, respectively.  $\nabla$  represents the first-order difference operation in the horizontal and vertical directions.

(4) Illumination Smoothness Loss: According to Retinex theory, the illumination should be smooth, allowing the details of the image to be preserved in the reflectance map. Using *S* to represent  $\frac{1}{L}$ , the illumination smoothness loss  $L_4$  can be defined as follows:

$$L_4 = \|\nabla S\|_2^2 + \|\nabla S'\|_2^2 \tag{14}$$

(5) The overall loss of the network can be defined as follows:

$$L = L_1 + L_2 + L_3 + \alpha L_4 \tag{15}$$

where  $\alpha$  represents the weight of the illumination smoothness loss.

# 3. Experiments

3.1. Experimental Settings

- Dataset. The low-light image data used in this study were collected from intelligent mining system surveillance images of Xieqiao Mine in Huainan City. The dataset consists of 1100 pairs of low/normal light images. Additionally, 100 underground mining images were collected from Baidu and Google. All images were resized to 512 × 512 and named as the MINE dataset. It is worth mentioning that we selected only 1000 low-light images as the training set and 200 images as the test set.
- **Training.** The algorithm presented in this paper was implemented on a PC equipped with a CPU (Intel(R) Core(TM) i7-12700K) and a GPU (3090TI), using the PyTorch framework to construct the network model. The batch size was set to 8. We utilized the Adam optimizer with default parameters and set the learning rate to 10<sup>-4</sup>. Each input was subjected to brightness perturbation once, with a weight of 10. We trained the proposed method for 500 epochs and conducted evaluations every 50 epochs. During each evaluation, we assessed the model's performance using a validation set and recorded the model that achieved the best performance on the validation set. This best-performing model was selected as our final model and utilized for subsequent experiments and applications. This training and evaluation set, enhancing the robustness and generalization capability of our method [36].
- Evaluation Metrics. For the real-world paired datasets tested in our study, we recommend utilizing two full-reference distortion metrics: peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) [57]. PSNR and SSIM are two widely used metrics for evaluating the quality of image reconstruction or enhancement algorithms. PSNR measures the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of the signal. It quantifies the amount of noise or distortion present in an image by computing the mean squared error (MSE) between the original image and the reconstructed image. The PSNR value is typically expressed in decibels (dB), with higher values indicating better image quality. SSIM, on the other hand, assesses the structural similarity between two images by considering their luminance, contrast, and structure. It compares local image patches and computes a similarity index based on the mean, variance, and covariance of these patches. The SSIM index ranges from 0 to 1, where a value closer to 1 indicates a higher structural similarity and better image quality.
- Comparison Methods. To validate the effectiveness of the proposed method, we compared it with state-of-the-art (SOTA) methods from recent years, including RetinexNet [19], Zero-DCE [36], EnlightenGAN [38], SCI [47], URetinex-Net [25], SNRNet [32], WCDM [33], and UHDFour [34].

# 3.2. Results

# 3.2.1. Visual Comparison

From Figures 7 and 8, it can be observed that the original images were generally dark, with low illumination, and unclear image details. RetinexNet could enhance the brightness of the images but resulted in significant color distortion, as shown in Figures 7b and 8b. The Zero-DCE algorithm enhanced the overall brightness of the images, but the brightness became excessively strong, causing overexposure. EnlightenGAN and SCI could improve the global brightness of the images, but they resulted in a significant loss of image information. The URetinex-Net, SNRNet, WCDM, and UHDFour algorithms effectively enhanced the brightness and contrast of the images, but they also led to considerable loss of details. In comparison, our proposed method exhibited superior enhancement results. It not only effectively enhanced the overall brightness of the images but also improved the brightness of darker local regions, highlighting more image details, features, and edge information. Furthermore, our method effectively mitigated color distortion issues that can occur during the image enhancement process, preserving the original color of the images.



(f)URetinex-Net

(g)SNRNet

(h)WCDM

(i)UHDFour

(j)Ours

**Figure 7.** Visual comparisons of state-of-the-art low-light image enhancement methods on Scenario S1 low-light images from the MINE dataset are presented below. Please zoom in for better detail.



(f)URetinex-Net

(g)SNRNet

(h)WCDM (i)UHDFour

(j)Ours

**Figure 8.** Visual comparisons of state-of-the-art low-light image enhancement methods on Scenario S2 low-light images from the MINE dataset are presented below. Please zoom in for better detail.

To evaluate the practicality of the algorithm, we applied it to the monitoring system of Xieqiao Mine in Huainan City and conducted enhancement comparison experiments on real-time images from two sites, namely T1 and T2. The T1 image contained only static scenes, allowing us to assess the effectiveness of detail and texture enhancement. The T2 image included human figures, providing a basis for evaluating the model's enhancedetection system in the mine. The experimental results are shown in Figures 9 and 10. 923-05-29 14:14

ment performance on individuals and establishing the foundation for a real-time person



Figures 9 and 10 demonstrate that the other advanced algorithms exhibited halo artifacts, shadows, and color distortion, resulting in incomplete image information and significant loss of fine details. In comparison, our algorithm suppressed the overexposed

areas, leading to a noticeable reduction in brightness of individuals, while enhancing the brightness in darker regions. This enhanced the distinction between individuals and the environment, resulting in a visually pleasing effect. Thus, our algorithm effectively achieved low-light image enhancement in underground environments.

# 3.2.2. Quantitative Comparison

Table 1 presents a comparison of the PSNR and SSIM values for the enhancement results of our network and the other state-of-the-art methods on the S1, S2, T1, and T2 images. It can be observed that our proposed model achieved the highest PSNR and SSIM values. This indicates that our method effectively enhanced the quality and fidelity of low-light images in underground environments.

**Table 1.** Quantitative evaluation of different methods on S1 , S2, T1, and T2. The best results are highlighted in bold black.

Mathada	Reference	<b>S</b> 1		S2		T1		T2	
Methous		PSNR↑	<b>SSIM</b> ↑	<b>PSNR</b> ↑	<b>SSIM</b> ↑	<b>PSNR</b> ↑	<b>SSIM</b> ↑	<b>PSNR</b> ↑	<b>SSIM</b> ↑
RetinexNet	BMVC 18 [19]	13.69	0.404	12.58	0.457	13.84	0.458	12.87	0.421
Zero-DCE	CVPR 20 [36]	18.98	0.561	15.87	0.512	17.45	0.461	16.34	0.436
EnlightenGAN	TIP 21 [38]	17.11	0.681	16.84	0.534	18.36	0.532	17.32	0.597
SCI	CVPR 22 [47]	19.12	0.654	18.94	0.686	17.48	0.674	18.59	0.648
URetinex-Net	CVPR 22 [25]	18.46	0.674	18.13	0.699	17.38	0.668	18.64	0.687
SNRNet	CVPR 22 [32]	17.38	0.642	18.21	0.712	17.22	0.674	17.46	0.672
WCDM	TOG 23 [33]	20.07	0.718	21.13	0.708	19.35	0.716	18.98	0.694
UHDFour	ICLR [34]	19.51	0.695	20.05	0.722	18.47	0.703	17.21	0.683
Ours	-	22.94	0.813	23.46	0.804	22.84	0.708	21.21	0.723

We conducted an evaluation of images from four distinct scenarios. From the results presented in Table 1, it can be concluded that our method achieved an approximately 14.3% improvement in PSNR and an approximately 13.2% improvement in SSIM compared to the WCDM competitor in the S1 scenario. In the S2 scenario, our method demonstrated approximately 11.1% and 11.4% improvements in PSNR and SSIM, respectively, relative to both the WCDM and UHDFour competitors. Regarding the T1 scenario, our method exhibited an approximately 18.0% enhancement in PSNR compared to the WCDM competitor, along with an approximately 1.1% improvement in SSIM. In the T2 scenario, our method achieved approximately 11.8% and 4.2% improvements in PSNR and SSIM, respectively, compared to the WCDM competitor. In summary, our method showcased excellent performance across diverse scenarios, demonstrating significant improvements over the competitors in multiple evaluation metrics. These findings underscore the substantial effectiveness and performance of our approach in the domain of image enhancement.

To enhance the authenticity of the experiments, we selected 200 images from the MINE dataset and performed a comparative enhancement using our algorithm and other advanced algorithms. The average PSNR and average SSIM were calculated, and the results are presented in Table 2.

Upon analyzing the Table 2, our method achieved the highest PSNR value of 22.18, which represents a significant improvement compared to the other methods. Specifically, it outperformed the next best method, WCDM, by 10.3%. In terms of SSIM, our method also performed remarkably well, with a value of 0.807, surpassing all other methods. Notably, it exhibited an improvement of 19.5% over the second-best performer, WCDM. These results demonstrate that our proposed method significantly enhanced the quality and similarity of the images, surpassing the performance of the other algorithms by a considerable margin.

Methods	Reference	<b>PSNR</b> ↑	SSIM↑
RetinexNet	BMVC 18 [19]	13.58	0.457
Zero-DCE	CVPR 20 [36]	17.99	0.558
EnlightenGAN	TIP 21 [38]	17.05	0.649
SCI	CVPR 22 [47]	19.41	0.671
URetinex-Net	CVPR 22 [25]	18.59	0.659
SNRNet	CVPR 22 [32]	18.01	0.638
WCDM	TOG 23 [33]	20.11	0.709
UHDFour	ICLR 23 [34]	19.68	0.687
Ours	-	22.18	0.807

**Table 2.** Quantitative evaluation of the different methods on MINE. The best results are highlighted in bold black.

# 3.3. Ablation Study

To validate the effectiveness of the proposed improvement method, comparative analysis was conducted through ablation experiments. The experimental results are presented in Figure 11. In group (a), low-light images were used, while group (b) represents the enhanced results using our algorithm. For groups (c), (d), (e), and (f), the experiments involved keeping the modules in the network unchanged while conducting ablation experiments on each loss function. Similarly, for groups (g), (h), and (i), the experiments involved keeping the loss functions unchanged while conducting ablation experiments on various modules within the network. The test quality indicators were selected from the MINE dataset, which consists of low-light images captured in underground mines. The comparative experimental results are shown in Figure 11, and the objective metric evaluations are presented in Table 3.

Methods **PSNR**<sup>↑</sup> SSIM<sup>↑</sup> Ours 0.723 21.46  $w/o L_1$ 18.290.698  $w/o L_2$ 15.53 0.558  $w/o L_3$ 18.42 0.662 $w/o L_4$ 17.56 0.597 w/o ResNeSt Block 20.01 0.701 w/o Attention 18.52 0.687 w/o Interference 19.31 0.679

Table 3. Quantitative results of ablation studies. The best results are highlighted in bold black.

From Figure 11c, it can be observed that the removal of the exposure control loss  $L_1$  resulted in an overall darkening of the image. Figure 11d demonstrates that in the experiment without the reflectance consistency loss  $L_2$ , the images exhibited abnormal exposure. Despite the introduced interference, the network fundamentally mapped from image to image, rather than image to reflectance, due to the absence of reflectance consistency loss. Figure 11e reveals that the removal of the structural loss  $L_3$  introduced artifacts and compromised some texture details. Figure 11f shows that in the absence of the illumination smoothness loss  $L_4$ , the image contrast became excessively high. From Figure 11g, it is evident that removing the ResNeSt block lad to a significant loss of image details, due to the decreased feature extraction capability of the network. Furthermore, Figure 11h demonstrates that removing the attention module resulted in overexposure and partial blurring of details. In Figure 11i, the removal of the random brightness perturbation module caused excessive enhancement, leading to color distortions.



(a)Input

(b)Ours





(d)w/o L<sub>2</sub>



(e)w/o L<sub>3</sub>









(h)w/o Attention

(i)w/o Interference

Figure 11. Visual comparison of the effectiveness of the ablated modules in the conducted research.

Based on the comparison of PSNR values in Table 1, it is evident that the "Ours" method outperformed the "w/o" variants in terms of image quality. The highest PSNR value achieved was 21.46, attributed to the "Ours" method, indicating its superior performance in preserving image details and reducing noise compared to the other variants. In terms of SSIM, the "Ours" method generally exhibited better performance than the "w/o" variants, although the differences were not significantly pronounced compared to the PSNR values. The SSIM value for the "Ours" method was 0.723, indicating a higher structural similarity between the enhanced and original images. Specific evaluations were also conducted for particular method variants. It can be observed that the exclusion of certain components, such as the "ResNeSt Block" or "Attention", led to a decrease in both PSNR and SSIM values. This highlights the crucial role of these components in the overall performance and effectiveness of the method.

These research findings demonstrate the high efficacy of the proposed method in addressing image quality issues under dust and low-illumination conditions in coal mine environments. This has significant implications for the safety, production efficiency, and environmental monitoring of coal mine workers. Overall, this study's contribution lies in its innovative approach to improving image quality in coal mine environments. The proposed method has the potential to enhance the safety and productivity of coal mining operations and is of great importance for the field of image processing in the context of coal mines.

# 4. Conclusions

This study presented a coal mine image enhancement method based on the HSV color space, to address the impact of dust and low-illumination conditions on video image enhancement in coal mine environments. The proposed method combines the characteristics of the HSV color space and applies Retinex theory to the brightness (*V* channel). By designing a reflection estimation network based on U-Net and incorporating ResNeSt blocks and multi-scale channel-wise pixel attention modules, the accuracy and generalization ability of the algorithm were improved. Experimental results demonstrated the effectiveness of our method in addressing image quality issues caused by dust and low-illumination conditions in coal mine environments. Our method achieved the highest PSNR value of 22.18, representing a significant improvement over the other methods, with a 10.3% enhancement compared to the second-best WCDM method. Additionally, our method performed remarkably well in terms of SSIM, achieving a value of 0.807, surpassing all other methods, and exhibiting a noteworthy improvement of 19.5% over the second-best performer, WCDM. In conclusion, our proposed method effectively addresses image quality issues caused by dust and low-illumination conditions in coal mine environments.

While this study makes significant contributions, there are also some limitations. One limitation was the dataset used for evaluation, which may have limited the diversity of coal mine image samples. Future research should consider expanding the dataset to include a greater variety of coal mine types and image variations under different environmental conditions, to increase the representativeness of the experiments. Another limitation was that the proposed method was specifically designed and evaluated for coal mine environments, and its applicability in other domains or general image enhancement tasks remains to be explored. Future research could investigate the transferability and adaptability of this method to different industries or broader image-processing applications. Lastly, although the proposed method achieved promising results, there is still room for improvement in algorithm efficiency and computational complexity. Future research could focus on optimizing the computational aspects to make the method more practical for real-time or large-scale applications. Therefore, future research should consider further exploration in areas such as multi-modal image enhancement, data augmentation, and synthesis, as well as joint optimization and task relevance.

In summary, this study proposed an unsupervised coal mine image enhancement method in the HSV color space, which showed promising results in low-illumination coal mine image applications. Furthermore, this method can also be applied to other industrial domains, considering that many industrial environments, similarly to coal mine settings, often suffer from low-lighting conditions and dust-related issues. By improving the quality of industrial video images, this method can enhance the effectiveness of visualization, monitoring, and detection systems. Future research could focus on further refining and expanding this method to enhance its applicability and effectiveness, as well as exploring its integration with other tasks, thus providing a more comprehensive solution for coal mine image processing.

**Author Contributions:** Conceptualization, L.W.; methodology, C.W.; data curation, D.W.; formal analysis, L.W. and W.D.; funding acquisition, K.H.; investigation, L.W.; methodology, L.W.; validation, D.W.; writing—original draft, C.W., L.W. and D.W.; writing—review and editing, C.W., K.H., D.W. and L.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the State Key Laboratory Open Fund of Mining Response and Disaster Prevention, Control in Deep Coal Mines (grant number SKLMRDPC21KF23) and the Training Programme Foundation for the Talents by Universities of Anhui Province (grant number gxyq2022068), The Foundation of Excellent Middle and Young Teachers Cultivation Action Project of Anhui Province (No.YQYB2023031), Key scientific research projects of colleges and universities in Henan Province (21A590001) and Key scientific research projects in colleges and universities of Anhui Provincial Department of Education (2023AH051547).

**Data Availability Statement:** The data presented in this study is available on request from the corresponding authors, and the dataset was jointly completed by the team, so the data is not publicly available.

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

#### References

- 1. Ding, E.J.; Yu, X.; Xia, B. Development of mine iforatization and key technologies of ielligent mines. *J. China Coal Soc.* 2022, 47, 564–578.
- 2. Chen, W.; Ren, P.; Tian, Z.J. Unsupervised mine personnel tracking based on attention mechanism. *J. China Coal Soc.* 2021, 46, 601–608.
- 3. Han, J.H.; Wei, X.; Lu, Y. Driverless technology of underground locomotive in coal mine. J. China Coal Soc. 2020, 45, 2104–2115.
- 4. Yang, X.; Chen, W.; Ren, P. Coal mine monitoring image semantic segmentation based on domain adaptation. *J. China Coal Soc.* **2021**, *46*, 3386–3396.
- 5. Shan, P.F.; Sun, H.Q.; Lai, X.P. Identifica-tion method on mixed and release state of coal-gangue masses of fully mechanized caving based on improved Faster R-CNN. *J. China Coal Soc.* **2022**, 47, 1382–1394.
- Cheng, D.Q.; Qian, J.S.; Guo, X.G. Review On key technologies of AI recognition for videos in coal mine. *Coal Sci. Technol.* 2023, 51, 349–365.
- 7. Wang, H.J.; Wang, H.L.Status and prospect of ielligent key technologies of belt conveyor. Coal Sci. Technol. 2022, 50, 225–239.
- 8. Xu, Y.; Wen, J.; Fei, L.; Zhang, Z. Review of Video and Image Defogging Algorithms and Related Studies on Image Restoration and Enhancement. *IEEE Access* 2015, *4*, 165–188. [CrossRef]
- 9. Zimmerman, J.; Pizer, S.; Staab, E.; Perry, J.; Mccartney, W.; Brenton, B. An evaluation of the effectiveness of adaptive histogram equalization for contrast enhancement. *IEEE Trans. Med. Imaging* **2015**, *1988*, 304–312. [CrossRef]
- 10. Lee, S.; Yun, S.; Nam, J.-H.; Won, C.S.; Jung, S.-W. A review on dark channel prior based image dehazing algorithms. *EURASIP J. Image Video Process.* **2016**, *4*, 1–23. [CrossRef]
- 11. Land, E.H.; McCann, J.J. Lightness and retinex theory. Josa 1971, 61, 1–11. [CrossRef] [PubMed]
- 12. Tian, F.; Chen, T.; Zhang, J. Research on Improved Retinex-Based Image Enhancement Method for Mine Monitoring. *Appl. Sci.* **2023**, *13*, 2672. [CrossRef]
- 13. Tian, F.; Wang, M.; Liu, X. Low-Light Mine Image Enhancement Algorithm Based on Improved Retinex. *Appl. Sci.* 2024, *14*, 2213. [CrossRef]
- 14. Liu, L.; Wang, H. Image enhancement using a nonlinear method with an improved Single-Scale Retinex algorithm. In Proceedings of the 2011 International Conference on Electronics, Communications and Control (ICECC), Ningbo, China, 9–11 September 2011; pp. 2086–2089.
- 15. Rahman, Z.-U.; Jobson, D.J.; Woodell, G.A. Multi-scale retinex for color image enhancement. In Proceedings of the 3rd IEEE International Conference on Image Processing, Lausanne, Switzerland, 19 September 1996; pp. 1003–1006.
- 16. Jiang, X.F.; Wang, G.; Shen, W.-M. A method of color image enhancement using color advanced retinex. *J. Optoelectron. Laser* **2008**, *19*, 1402–1404.
- 17. Lore, K.G.; Akintayo, A.; Sarkar, S. LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognit*. **2017**, *61*, 650–662. [CrossRef]
- 18. Zhang, Y.; Guo, X.; Ma, J.; Liu, W.; Zhang, J. Beyond brightening low-light images. *Int. J. Comput. Vis.* **2021**, 129, 1013–1037. [CrossRef]
- 19. Wei, C.; Wang, W.; Yang, W.; Liu, J. Deep retinex decomposition for low-light enhancement. arXiv 2018, arXiv:1808.04560.
- 20. Jiang, K.; Wang, Z.; Wang, Z.; Chen, C.; Yi, P.; Lu, T.; Lin, C.-W. Degrade is upgrade: Learning degradation for low-light image enhancement. In Proceedings of the AAAI Conference on Artificial Intelligence, Online, 22 February–1 March 2022; pp. 1078–1086.
- 21. Xu, K.; Chen, H.; Tan, X.; Chen, Y.; Jin, Y.; Kan, Y.; Zhu, C. HFMNet: Hierarchical feature mining network for low-light image enhancement. *IEEE Trans. Instrum. Meas.* 2022, 71, 1–14. [CrossRef]
- 22. Lu, Y.; Guo, Y.; Liu, R.W.; Ren, W. MTRBNet: Multi-branch topology residual block-based network for low-light enhancement. *IEEE Signal Process. Lett.* **2022**, *29*, 1127–1131. [CrossRef]
- 23. Fan, G.-D.; Fan, B.; Gan, M.; Chen, G.-Y.; Chen, C.P. Multiscale low-light image enhancement network with illumination constraint. *IEEE Trans. Circuits Syst. Video Technol.* 2022, *32*, 7403–7417. [CrossRef]
- 24. Gao, S.-H.; Cheng, M.-M.; Zhao, K.; Zhang, X.-Y.; Yang, M.-H.; Torr, P. Res2net: A new multi-scale backbone architecture. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *43*, 652–662. [CrossRef] [PubMed]
- 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 IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA,18–24 June 2022; Volume 129, pp. 5901–5910.
- 26. Lu, K.; Zhang, L. TBEFN: A two-branch exposure-fusion network for low-light image enhancement. *IEEE Trans. Multimed.* **2020**, 23, 4093–4105. [CrossRef]

- Lu, H.; Gong, J.; Liu, Z.; Lan, R.; Pan, X. FDMLNet: A Frequency-Division and Multiscale Learning Network for Enhancing Low-Light Image. Sensors 2022, 22, 8244. [CrossRef]
- Xu, K.; Chen, H.; Xu, C.; Jin, Y.; Zhu, C. Structure-texture aware network for low-light image enhancement. *IEEE Trans. Circuits* Syst. Video Technol. 2022, 32, 4983–4996. [CrossRef]
- Garg, A.; Pan, X.-W.; Dung, L.-R. LiCENt: Low-light image enhancement using the light channel of HSL. *IEEE Access* 2022, 10, 33547–33560. [CrossRef]
- Chen, J.; Lian, Q.; Zhang, X.; Zhang, D.; Yang, Y. Hcsam-Net: Multistage Network with a Hybrid of Convolution and Self-Attention Mechanism for Low-Light Image Enhancement. 2022. Available online: https://ssrn.com/abstract=4237486 (accessed on 2 January 2024).
- Yang, Y.; Hu, W.; Huang, S.; Tu, W.; Wan, W. Low-light image enhancement network based on multi-stream information supplement. *Multidimens. Syst. Signal Process.* 2022, 33, 711–723. [CrossRef]
- Xu, X.; Wang, R.; Fu, C.-W.; Jia, J. Snr-aware low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–20 June 2022; pp. 17714–17724.
- Jiang, H.; Luo, A.; Fan, H.; Han, S.; Liu, S. Low-light image enhancement with wavelet-based diffusion models. ACM Trans. Graph. (TOG) 2023, 42, 1–14. [CrossRef]
- 34. Li, C.; Guo, C.-L.; Zhou, M.; Liang, Z.; Zhou, S.; Feng, R.; Loy, C.C. Embedding fourier for ultra-high-definition low-light image enhancement. *arXiv* 2023, arXiv:2302.11831.
- 35. Jin, Y.; Yang, W.; Tan, R.T. Unsupervised night image enhancement: When layer decomposition meets light-effects suppression. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; pp. 404–421.
- Guo, C.; Li, C.; Guo, J.; Loy, C.C.; Hou, J.; Kwong, S.; Cong, R. Zero-reference deep curve estimation for low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 1780–1789.
- Li, C.; Guo, C.; Loy, C.C. Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* 2021, 44, 4225–4238. [CrossRef]
- Jiang, Y.; Gong, X.; Liu, D.; Cheng, Y.; Fang, C.; Shen, X.; Yang, J.; Zhou, P.; Wang, Z. Enlightengan: Deep light enhancement without paired supervision. *IEEE Trans. Image Process.* 2021, 30, 2340–2349. [CrossRef]
- Fu, Y.; Hong, Y.; Chen, L.; You, S. LE-GAN: Unsupervised low-light image enhancement network using attention module and identity invariant loss. *Knowl.-Based Syst.* 2022, 240, 108010. [CrossRef]
- 40. Ni, Z.; Yang, W.; Wang, H.; Wang, S.; Ma, L.; Kwong, S. Cycle-interactive generative adversarial network for robust unsupervised low-light enhancement. In Proceedings of the 30th ACM International Conference on Multimedia, Lisboa, Portugal, 10–14 October 2022; pp. 1484–1492.
- 41. Wang, R.; Jiang, B.; Yang, C.; Li, Q.; Zhang, B. MAGAN: Unsupervised low-light image enhancement guided by mixed-attention. *Big Data Min. Anal.* 2022, *5*, 110–119. [CrossRef]
- 42. Qiao, J.; Wang, X.; Chen, J.; Jian, M. Low-light image enhancement with an anti-attention block-based generative adversarial network. *Electronics* **2022**, *11*, 1627. [CrossRef]
- Xu, W.; Chen, X.; Guo, H.; Huang, X.; Liu, W. Unsupervised image restoration with quality-task-perception loss. *IEEE Trans. Circuits Syst. Video Technol.* 2022, 32, 5736–5747. [CrossRef]
- 44. Jiang, Q.; Mao, Y.; Cong, R.; Ren, W.; Huang, C.; Shao, F. Unsupervised decomposition and correction network for low-light image enhancement. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 19440–19455. [CrossRef]
- 45. Zhu, J.-Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.
- 46. Bhattacharya, J.; Modi, S.; Gregorat, L.; Ramponi, G. D2bgan: A dark to bright image conversion model for quality enhancement and analysis tasks without paired supervision. *IEEE Access* **2022**, *10*, 57942–57961. [CrossRef]
- 47. Ma, L.; Ma, T.; Liu, R.; Fan, X.; Luo, Z. Toward fast, flexible, and robust low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 5637–5646.
- Liu, S.; Long, W.; He, L.; Li, Y.; Ding, W. Retinex-based fast algorithm for low-light image enhancement. *Entropy* 2021, 23, 746. [CrossRef] [PubMed]
- 49. Peng, L.; Huang, Y.; Yao, K. Multi-algorithm fusion of RGB and HSV color spaces for image enhancement. In Proceedings of the 37th Chinese Control Conference (CCC), Wuhan, China, 25–27 July 2018; pp. 9584–9589.
- Zhang, Y.; Zhang, J.; Guo, X. Kindling the darkness: A practical low-light image enhancer. In Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 21–25 October 2019; pp. 1632–1640.
- 51. Zotin, A. Fast algorithm of image enhancement based on multi-scale retinex. Procedia Comput. Sci. 2018, 131, 6–14. [CrossRef]
- Wang, R.; Zhang, Q.; Fu, C.-W.; Shen, X.; Zheng, W.-S.; Jia, J. Underexposed photo enhancement using deep illumination estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 6849–6857.
- Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference (Part III 18), Munich, Germany, 5–9 October 2015; pp. 234–241.

- 54. Zhang, H.; Wu, C.; Zhang, Z.; Zhu, Y.; Lin, H.; Zhang, Z.; Sun, Y.; He, T.; Mueller, J.; Manmatha, R. Resnest: Split-attention networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 2736–2746.
- 55. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
- Li, X.; Wang, W.; Hu, X.; Yang, J. Selective kernel networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 510–519.
- 57. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* 2004, 13, 600–612. [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.