IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 12 March 2024, accepted 15 May 2024, date of publication 20 May 2024, date of current version 29 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3403452

APPLIED RESEARCH

Enhancement of Mine Images Based on HSV Color Space

CHANGLIN WU[1], DANDAN WANG[1,2], AND KAIFENG HUANG[1]
[1]School of Mechanical and Electrical Engineering, Huainan Normal University, Huainan 232001, China
[2]College of Electronic Information and Electrical Engineering, Anyang Institute of Technology, Anyang, Henan 455000, China

Corresponding author: Dandan Wang (lansejingling1988@126.com)

This work was supported in part by the Key Research Projects of Higher Education Institutions in Henan Province under Grant 21A590001, in part by Anhui Provincial Department of Education's Key Scientific Research Projects in Universities under Grant 2023AH051547, in part by the Research on Intelligent Diagnosis Technology of Safety Monitoring Sensor Based on Deep Learning Algorithm, in part by the Open Fund of State Key Laboratory for Mining Response and Disaster Prevention and Control of Deep Coal Mines under Grant SKLMRDPC21KF23, and in part by Anhui Provincial Department of Education for the Outstanding Youth Talent Support Program in Universities under Grant gxyq2022068.

ABSTRACT This study focuses on enhancing the quality of video images in coal mines under challenging conditions such as dust and low illumination. Existing algorithms often suffer from poor generalization and low accuracy. To address these limitations, we propose an unsupervised image enhancement method based on the HSV color space transformation, incorporating the Retinex theory into the luminance component (V channel). A disturbance technique is employed to perturb the luminance, and a reflectance estimation network based on U-Net is designed to ensure consistency between the original reflectance and the disturbed reflectance within the same scene. Additionally, residual multiscale and attention mechanism modules are introduced to improve accuracy while reducing the network's parameter count. The saturation component (S channel) is adaptively adjusted based on the correlation coefficient. The final enhanced image is obtained by recombining the original hue (H channel), luminance, and saturation before converting to the RGB color space. Importantly, our algorithm does not require training on normal light images. The experimental results indicate that our algorithm outperforms other state-of-the-art algorithms in terms of objective quality metrics, namely Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). Additionally, our algorithm exhibits superior performance in subjective visual analysis compared to the comparative algorithms, demonstrating its efficacy in improving the visual quality of low-light images in mining environments.

INDEX TERMS Coal mine images, Retinex, U-Net, random disturbance.

I. INTRODUCTION

With the continuous advancement of information construction in high-tech coal mines [1], traditional coal mining operations are gradually transitioning towards automatic unmanned mining. In this context, the intelligence level of digital video monitoring systems [2] in coal mines is also improving, providing reliable data support for downstream tasks such as target detection [3], [4], special area early warning [5], and target tracking [6]. However, underground coal mines face challenges such as high dust and fog concentrations, as well as low brightness from artificial light sources. These factors result in low illumination, color distortion, and loss of detail in the image data captured by underground camera devices. Therefore, there is a need to enhance low illumination images in coal mines [7] in order to improve image quality.

Traditional low-light image enhancement algorithms encompass histogram equalization [8], [9], dark channel prior enhancement [10], and Retinex algorithms [11], [12]. Among these, Retinex algorithms, including Single-Scale Retinex (SSR) [13], Multi-Scale Retinex (MSR) [14], and Color Restoration Retinex (MSRCR) [15], are widely applied.

Nevertheless, traditional methods often exhibit disadvantages such as high computational complexity and poor generalization performance. To address the limitations of

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

traditional enhancement methods, in recent years, numerous scholars have employed deep neural network methods for enhancing low-light images. Depending on the approach, these methods can be classified into supervised and unsupervised methods.

In the realm of supervised learning-based approaches, Zhang et al. [16] proposed KinD, a neural network designed for enhancing low-light images. The method effectively addresses significant visual issues and allows for flexible contrast adjustments, leading to improved image quality. Expanding on this work, Zhang et al. [17] introduced an enhanced version called KinD++, which incorporates an additional mapping function to better align the enhanced images with real-world conditions. In a different vein, Wang et al. [18] introduced a novel enhancement technique based on normalized flow. Through the utilization of a normalized flow model, they successfully learned the correspondence between low-light and normal-light images, resulting in improved contrast and effective noise reduction. Jiang et al. [19] introduced a novel Degradation-to-Refinement Generative Network (DRGN) that focuses on enhancing fine image details while maintaining naturalness in the visual enhancements. Xu et al. [20] proposed HFMNet, a hierarchical feature mining network that extracts lighting and edge features at different network layers. Lu et al. [21] presented a multi-branch topological residual network that utilizes topological residual blocks to enhance the network's learning capacity and improve image quality. Zhuang et al. [22] introduced a dual-branch dilated network with phase-aware Fourier convolution for low-light image enhancement. Fan et al. [23] developed an illumination-constrained multi-scale low-light image enhancement network, leveraging Res2Net [24] to extract deep multi-scale features and mitigate color distortion, resulting in more natural visual effects. Wu et al. [25] proposed Retinex Net, a deep unfolding network based on the Retinex theory, which combines an implicit prior regularization model with the Retinex theory to effectively suppress noise and preserve image details. Lu et al. [26] addressed these challenges by proposing a dual-branch exposure fusion network for low-light image enhancement. Furthermore, the Deep Lighting Network (DLN) [27] treats low-light image enhancement as a residue learning problem and incorporates the inverse projection module from the image super-resolution domain into low-light image enhancement for the first time.

However, acquiring fully paired datasets for low-light image enhancement presents significant challenges, which restrict the practical effectiveness of supervised learning models. To address this issue, an increasing number of unsupervised and zero-shot learning models have been proposed. Jin et al. [28] addressed the issues of excessive and insufficient enhancement in low-light image enhancement by introducing an unsupervised enhancement method. This method incorporates a layer decomposition network and a light effect suppression network to further suppress

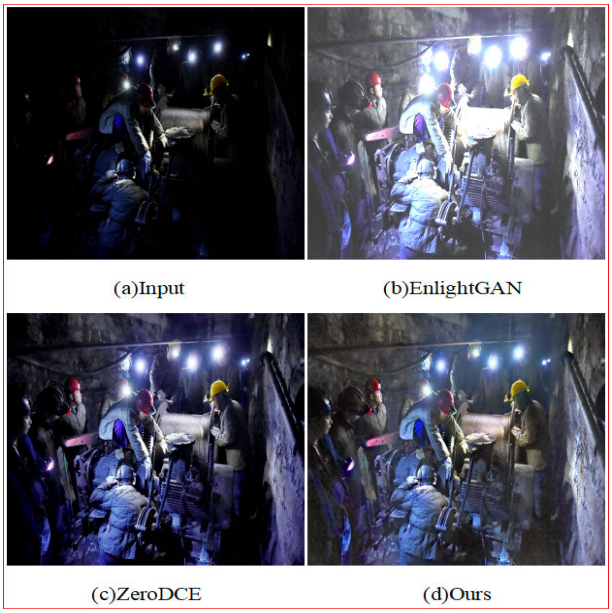

FIGURE 1. An example of mine low-light image enhanced by current unsupervised learning-based methods and the proposed method.

highlights and enhance dark regions, ensuring optimal contrast. Fu et al. [29] introduced LE-GAN, a novel unsupervised low-light image enhancement network based on generative adversarial networks trained using non-paired low-light images. Ni et al. [30] proposed CIGAN, a cyclic interactive generative adversarial network for unsupervised low-light image enhancement. This network comprises three components and incorporates attention mechanisms to improve image quality. Wang et al. [31] presented MAGAN, a mixed attention-guided generative adversarial network that employs a hybrid attention layer to capture the relationships between pixels and images, achieving image enhancement while reducing noise.

Qiao et al. [32] proposed a generative adversarial network (GAN) approach based on the reverse attention mechanism. Xu et al. [33] introduced a novel quality-aware loss to further enhance image quality. Jiang et al. [34] presented an unsupervised decomposition and correction network and incorporated a noise removal network to eliminate noise. Taking inspiration from CycleGAN [35], Bhattacharya et al. [36] proposed D2BGAN, an image enhancement model that effectively removes artifacts by combining geometric and illumination consistency, contextual loss, and multi-scale color, texture, and edge discriminators. EnlightenGAN [37] and Zero-DCE [38] offer two innovative approaches: the former utilizes an attention-guided U-Net [39] as the generator and employs a global-local discriminator, demonstrating the ability to perform unsupervised training using unpaired data and achieving impressive results and generalization. The latter transforms the low-light image enhancement problem into a curve estimation problem and iteratively enhances the

image using a network-simulated curve. It also meticulously designs a set of non-reference loss functions for unsupervised training. Li et al. [40] further introduced Zero DCE++, an accelerated and lightweight version of Zero-DCE. These deep convolutional neural network-based algorithms have shown significant advancements in performance compared to traditional methods. However, they still have limitations: as depicted in Figure 1, EnlightenGAN introduces color biases in the enhanced images, and Zero-DCE may result in blurry images with a lack of details and sharpness. The unsupervised methods have potential for improvement in terms of enhancement effectiveness.

In summary, both supervised and unsupervised methods exhibit certain limitations:

1)Supervised methods rely on paired datasets, which are often challenging to acquire. The scarcity of paired data hampers the training process and limits the availability of suitable datasets.

2)Supervised methods may lack generalization capabilities. Since they heavily rely on paired data for training, they tend to perform well on specific datasets but struggle when applied to different scenarios. The models may fail to effectively adapt to new and unseen data.

3)Unsupervised methods often face challenges in achieving optimal enhancement results. In complex lighting conditions, such as those encountered in underground mines, the lighting varies across different regions of captured images. This makes it difficult to strike a balance between preserving details in darker regions and avoiding overexposure in brighter areas, as observed in Figure 1. Consequently, there is a risk of losing information in overexposed regions while still failing to adequately enhance darker areas.

In view of the problems existing in the above algorithms, this paper proposes an unsupervised lightweight network model, which converts the downhole RGB image into HSV space and enhances it with Retinex theory. The main works are as follows:

1) an unsupervised neural network model that combines a U-Net network, residual multi-scale convolution, and attention mechanism is proposed. The model takes U-Net as the backbone network structure and consists of three parts: encoding path, attention module, and decoding path. Firstly, the image is sampled by residual multi-scale convolution and coding downsampling. at the same time, the characteristic tensor output of each level is copied to the decoder of the corresponding level through the attention module, and then the feature tensor obtained by downsampling is superimposed and upsampled. Finally, V-channel enhancement is realized through the loss function.

2) based on the improved U-Net model, a depth reflectivity estimation network is designed, and a loss function is designed to prevent the model from being over-fitted, and regularized by the consistency of luminance in the form of original picture and random disturbance.

3) A new random luminance disturbance method is designed, which is helpful in capturing the reflection changes that may occur in the actual scene under different illumination conditions. This method can improve the robustness and generalization of the model.

4) optimize the borehole wall image saturation and boundary blur after enhancement. While the luminance component is enhanced, the saturation component is also adaptively adjusted according to the correlation coefficient.

In conclusion, our proposed approach presents an unsupervised lightweight network model for enhancing Mine images. By incorporating U-Net, residual multi-scale convolution, attention mechanism, and Retinex theory, we demonstrate notable improvements in image quality and address specific challenges encountered in underground scenarios.

II. IMAGE ENHANCEMENT NETWORK MODEL BASED ON HSV SPATIAL TRANSFORMATION

A. OVERALL MODEL ARCHITECTURE

Aiming to address the limitations of existing mine image enhancement algorithms in terms of their poor enhancement effect and generalization ability, this paper proposes an image enhancement model in the HSV color space, as illustrated in Figure 2. The approach involves converting the mine image from the RGB color space to the HSV color space, extracting the hue (H), saturation (S), and luminance (V) components. The hue component (H) remains unchanged, while the Retinex theoretical model is combined with the luminance value to enhance the luminance component (V) using an improved U-Net network. To improve the generalization of the model, a random perturbation method is applied to the luminance component, yielding a perturbed component (V'). By enforcing consistency between the original reflectance and the randomly perturbed reflectance within the same scene, the model's ability to generalize is enhanced. Moreover, the random perturbation method allows the network model to achieve better image enhancement without relying on normal reference images. The enhanced luminance components (V_c and V_c') are obtained through the improved U-Net model while ensuring consistency with the perturbed luminance components (V_c and V_c'). Simultaneously, the saturation component is adaptively adjusted based on the correlation coefficient to obtain the corrected saturation component (S_c). Finally, the enhanced HSV results are converted back to the RGB color space to output the enhanced image. This comprehensive approach aims to overcome the limitations of existing algorithms by effectively enhancing the luminance, saturation, and overall quality of mine images.

B. HSV COLOR SPACE

In this article, we use HSV color space, HSV is a color system more commonly used in people's lives, it is more in line with the way people describe colors, that is, what color the object is, how dark the color is, how bright the color is. H is hue, or called hue, and it describes the color in degrees; S is saturation, that is, the depth of the color; and V is hue, which

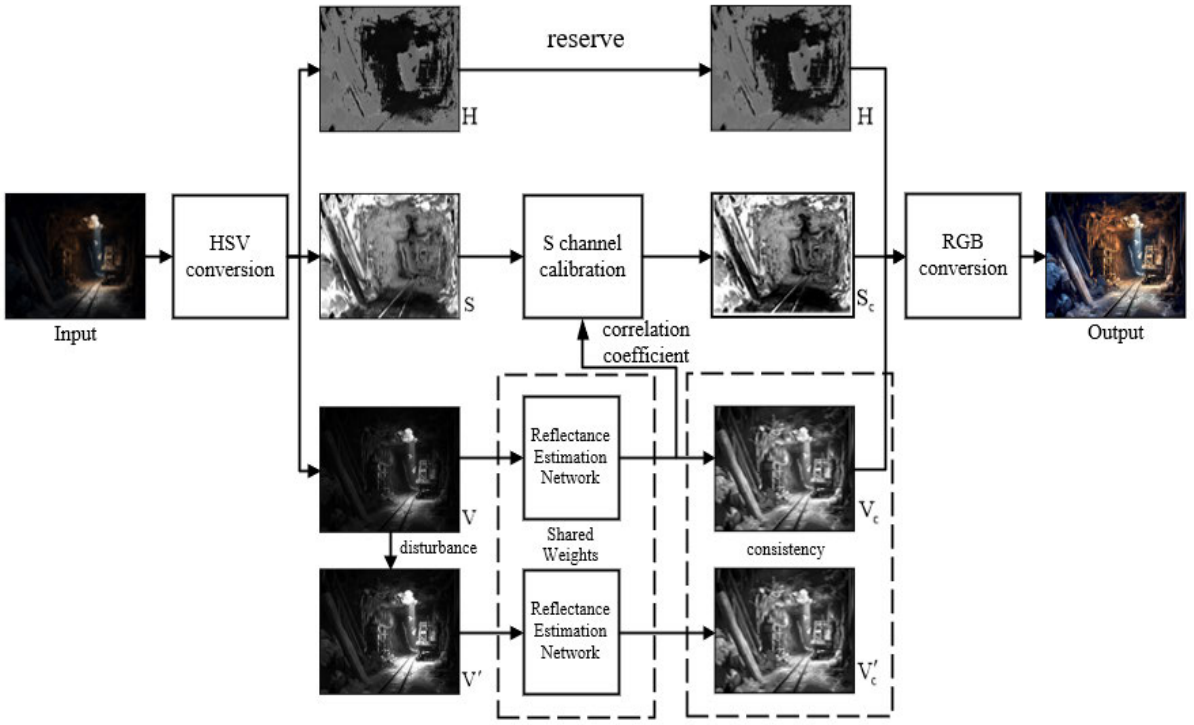

FIGURE 2. Overall model of mine image enhancement.

is the brightness of the color. The formula for converting an image from RGB color to HSV color space is as follows:

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

$$S = \begin{cases} 0 & C\max = 0 \\ \dfrac{\Delta}{C\max} & C\max \neq 0 \end{cases} \tag{2}$$

$$V = C\max \tag{3}$$

Among them, R', G',and B' normalize the values of R, G, and B to 0∼1.C max is the maximum value in R', G', B',and Δ is the difference between C max and C min. Compared to the RGB color space, the HSV color space offers several advantages in the mine image enhancement:

1) Intuitive expression of color information: Unlike the RGB channels, which do not accurately represent the color information of objects, the HSV color space provides a more intuitive representation of lightness, darkness, hue, and other color-related information. This facilitates better color comparison and analysis.

2) Preservation of color information: Converting the image to the HSV color space helps preserve the color information of the image during the enhancement process. It enables the effective recovery of color details in darker regions of the image and reduces the complexity of processing for the enhancement network.

3) Utilization of luminance channel (V): In an unevenly illuminated image, the luminance channel (V) in the HSV color space represents the largest channel. Each pixel value in this channel corresponds to the maximum value among the corresponding pixels in the three RGB channels. This characteristic provides a significant amount of information that can be leveraged to enhance mine images effectively.

C. RANDOM LUMINANCE DISTURBANCE

In order to capture the reflection changes that may occur in real-world scenes under varying lighting conditions, this paper introduces a random disturbance method for luminance. This method is implemented using a power function with a random exponent (the value to which the variable is raised) to generate an altered form of brightness within the same scene. It ensures that the disturbed form maintains the same reflectivity as the original form. Nonlinear functions, such as the one described in formula 4, offer several advantages:

1) Input and output range: The power function used in the random disturbance method has an input and output range of [0,1]. This range is chosen to avoid information loss caused by overflow truncation, ensuring that the enhanced image retains as much detail as possible.

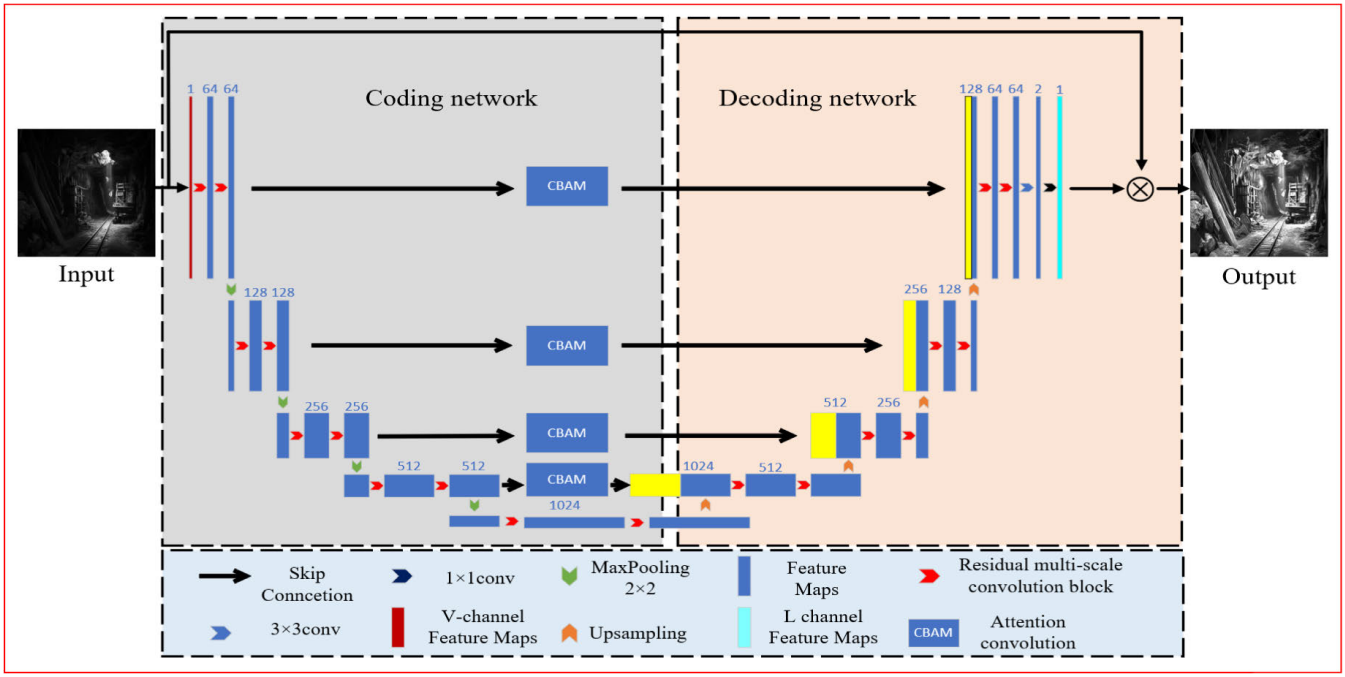

FIGURE 3. Improved U-Net reflectance estimation models.

2) Monotonically increasing function: The power function chosen for the disturbance method is monotonically increasing. This property helps maintain the consistency of the gradient direction between the original luminance form and the perturbed form. It ensures that the enhancement process preserves the overall structure and relationships within the image.

3) Random exponent for diversity: The use of a random exponent in the power function enhances the diversity of disturbances generated. By introducing randomness, the method can capture a wider range of reflection changes, making the model more robust and adaptable to different lighting conditions.

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

where $V(x)$ denotes pixel values of original V channel,$V'(x)$ is the disturbed V channel, d is a random value with its range determined by the average value of the original V channel. If the average value is less than 0.5, the range of d is [0,1], indicating that we use smaller d values to perturb the image. This is done to maintain subtle variations in the perturbation and avoid excessive impact on reflectivity. Conversely, if the average value of the original V channel is greater than or equal to 0.5, we choose a range of d as [1,4], meaning that we use larger d values to introduce more pronounced perturbation effects. This design allows us to adjust the intensity of the perturbation based on the average brightness level of the image, catering to the needs of different images. It is important to note that the purpose of this perturbation method is not to expand the training dataset, but rather to constrain the consistency of reflections. We aim to constrain the reflection characteristics in the image through perturbation, rather than increasing the training samples by increasing the data volume. This method effectively constrains reflection consistency and enhances the model's robustness to different reflection conditions.

D. REFLECTANCE ESTIMATION NETWORK BASED ON IMPROVED U-NET

Most networks based on the Retinex theory initially undergo training in the decomposition network to learn the process of separating the illumination component from the reflection component.However, this approach often introduces errors in the reconstruction of the image. The network generates an interfering luminance for a given disturbance, utilizing the same network architecture and shared parameters. According to the Retinex theory, an image can be divided into two distinct parts: the luminance and the reflection. The reflection component is expected to remain unchanged under different illumination conditions. Therefore, the aforementioned channel and its disturbance channel can be effectively decomposed into the following two parts.

$$\begin{cases} V = R * L \\ V' = R' * L' \end{cases} \tag{5}$$

where R is the reflection component and L is the illuminance component. If L is regarded as an intermediate variable for calculating R, the above formula can be changed to:

$$\begin{cases} R = V * \dfrac{1}{L} \\ R' = V' * \dfrac{1}{L'} \end{cases} \tag{6}$$

Because this does not have to calculate the difference between V and Runl, it can avoid losing information when reconstructing the image [41].

Based on the Retinex theory, we propose a U-Net-based network for reflectance estimation. In the same scene, the U-Net network demonstrates powerful capabilities in image enhancement tasks by estimating reflectance through enforcing consistency between the original reflectance and the reflectance perturbed by random noise. With its unique structure and design, the U-Net preserves image details, extracts and integrates context-aware features, resulting in enhanced images that are more accurate and natural. The adoption of skip connections allows information propagation between low-level and high-level features, enabling the U-Net to adapt to features at different scales and strike a balance between handling details and global information. Moreover, the flexibility and scalability of the U-Net make it applicable to various image enhancement tasks while maintaining efficient training and inference speeds.

Compared to other algorithms, the low-light image enhancement algorithm based on U-Net exhibits significant advantages. The classical U-Net consists of more than 31 million network parameters, making it time-consuming to fine-tune these parameters during network training using the gradient descent method. In this study, to cater to our specific application requirements and ensure better preservation of image information in the enhanced results, we propose a lightweight multi-scale residual attention U-Net model. This model combines the foundational structure of the traditional U-Net network with the benefits of inception multi-scale feature extraction, residual networks, and attention mechanisms. The model's architecture is depicted in Figure 3.

In the classical U-Net and its improved models, the convolution layers of the image are stacked, with smaller convolution kernels at the lower layers and larger convolution kernels at the upper layers, thereby increasing the network's depth. However, this approach not only escalates computational complexity but also tends to cause gradient vanishing. In our network, we replace the 3×3 convolution module in the U-Net model with residual multi-scale convolutions (indicated by the red arrow in Figure 3). This modification expands the network's width and enhances its adaptability to multi-scale targets. To address the issue of redundant information or noise being transmitted to the decoding layer simultaneously with the key detail features via the skip connections in the U-Net model, we introduce an attention mechanism into the network model (such as the CBAM module depicted in Figure 3).

E. RESIDUAL MULTI-SCALE CONVOLUTION BLOCK

At each level of the network illustrated in Figure 3, the feature tensor undergoes two consecutive residual multi-scale convolution blocks (indicated by the red arrow in Figure 3). The residual multi-scale convolution block is formed by combining the residual network and the inception module.

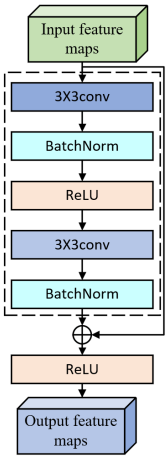

FIGURE 4. Residual structure.

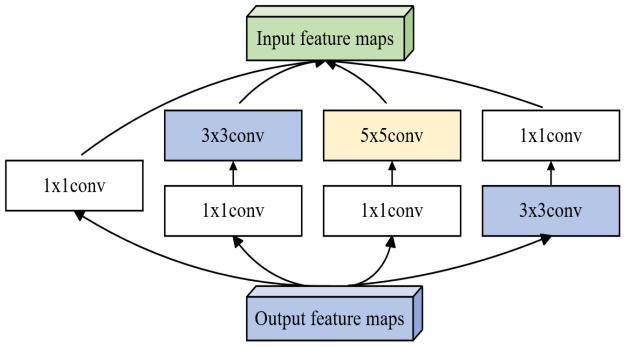

FIGURE 5. Inception structure.

The structure diagram of the residual (ResNet) network is presented in Figure 4. The residual module comprises two convolution layers, each consisting of a 3×3 convolution, batch normalization (BatchNormalization, BN) processing, and a modified rectified linear unit (ReLU). Following the two convolution layers, the input feature tensor is added to itself, and then passed through a ReLU activation function to yield the output feature tensor. The inclusion of a residual network significantly accelerates the convergence of the overall network. To enhance the local geometric details and generate more comprehensive high-level semantic information, the 3×3 convolution module in the residual structure is replaced by a multi-scale convolution Inception structure. This structure is illustrated in Figure 5. In this structure, the input feature tensor passes through four branches. In the first branch, the input tensor undergoes a 1×1 convolution. In the second branch, the output tensor is obtained by applying a 1×1 convolution in one branch and a 3×3 convolution in another branch. Similarly, in the third branch, the input tensor is sequentially convolved with a 1×1 convolution and a 5×5 convolution to obtain the output tensor. The fourth branch involves applying a 3×3 convolution and then a 1×1 convolution to the input tensor. The output tensors from the four branches are then

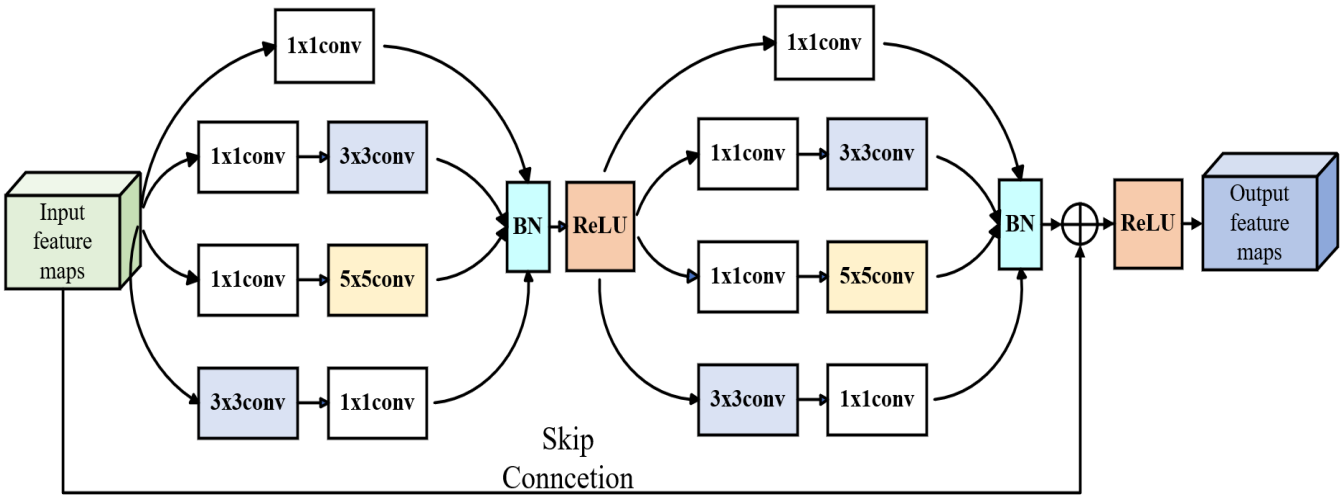

FIGURE 6. Residual multi-scale module.

concatenated along the channel dimension to obtain the final output tensor of the multi-scale convolution. This integration of the output tensors from different branches allows for the extraction of more diverse and comprehensive features.

Figure 6 illustrates the residual multi-scale module, which combines the residual network and the inception module. In this module, the 3×3 convolution in the residual network is replaced by the Inception structure, enabling deeper information mining and feature extraction. By incorporating the multi-scale convolution and the integration of features from different branches, the proposed network aims to capture both local details and high-level semantic information, resulting in enhanced performance for the given task.

F. ATTENTION MODULE

In the traditional U-Net network, the decoder is easy to loses some important details in the process of decoding, so the skip-connection connection is used in U-Net to map the feature information extracted from the encoder to the decoder through skip-connection. However, such a structure will cause some redundant information extracted in the encoder part to affect the feature fusion effect of the decoder part. In order to solve this problem, an attention mechanism can be used to solve this problem. The attention mechanism is added to the structure of the algorithm, which is similar to the human selective attention mechanism, which suppresses some irrelevant features in the learning process and strengthens the relevant features in the learning task. In this paper, we improve the skip-connection part of the U-Net network and introduce the CBMA module architecture. As shown in Figure 7, we can pay attention to the important features by using the attention mechanism, while suppressing the redundant features. The process of extracting information features by convolution operation is as follows: firstly, the module applies the channel attention

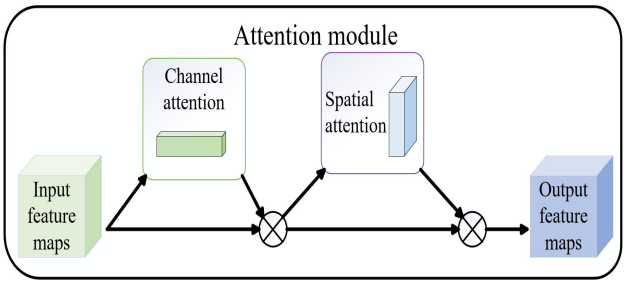

FIGURE 7. CBAM module.

module to learn the local feature information on the channel. Secondly, the spatial attention module is applied to learn the location feature information in space, so as to help the information flow in the network more efficiently. Given the feature map $F \in R^{C\times H\times W}$ as input, CBAM will calculate one-dimensional channel attention map $M_C \in R^{C\times 1\times 1}$ and two-dimensional spatial attention map $M_s \in R^{1\times W\times H}$ in turn, and the whole attention process can be summarized as follows:

$$F' = M_c(F) \otimes F \tag{7}$$

$$F'' = M_s(F') \otimes F' \tag{8}$$

In this context, let $\otimes$ denote element-wise multiplication. Firstly, the input feature map is multiplied with the channel attention map to obtain F'. Subsequently, the spatial attention map of F' is computed, and finally, the two are multiplied together to yield F''. Here, F'' represents the refined feature map.

III. DESIGN OF LOSS FUNCTION

To enhance low-light images in underground coal mines, we employed four sets of loss functions for model optimization: exposure loss function, reflectance consistency loss

function, spatial structure loss function, and illumination smoothness loss function. These loss functions were designed to address specific aspects of the image enhancement process, ensuring improved visibility and quality in challenging lighting conditions.

1) consistent loss of reflectivity. According to Retinex theory, the reflectivity of V and V' should be consistent, so the consistent loss of reflectivity L_1 can be defined as:

$$L_1 = \left\| R - R' \right\|_2^2 \tag{9}$$

where R and R are the reflectivity generated by V and V' respectively, and $\|\|_2$ is the norm of L_1.

2) exposure loss. To control the brightness of the generated reflectivity map, the difference between the average brightness of the reflectivity and the exposure value of a given normal image is calculated. If E is 0.7, the exposure loss L_2 can be defined as:

$$L_2 = \|R - E\|_2^2 \tag{10}$$

3)Space structure loss. To ensure that the spatial structure of the input image can be retained on the reflectivity map, the loss of spatial structure is introduced, which calculates the difference between the horizontal and vertical gradients of each pixel between the input image and the reflectivity map. The loss of spatial structure L_3 can be defined as:

$$L_3 = \|\nabla R_m - \nabla V_m\|_2^2 \tag{11}$$

R_m and V_m represent the generated reflectance and the average pooling result of the input, respectively. ∇ represents the first-order difference operation in the horizontal and vertical directions.

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

$$L_4 = \|\nabla S\|_2^2 + \left\|\nabla S'\right\|_2^2 \tag{12}$$

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

$$L = L_1 + L_2 + L_3 + \alpha L_4 \tag{13}$$

where α is the weight of the illumination smoothness loss.

IV. S CHANNEL ENHANCEMENT DESIGN

After enhancing the luminance component (V), the enhancement of luminance can also affect the saturation component (S). Therefore, it is necessary to correct the saturation component S. Common methods for saturation enhancement include linear stretching and histogram equalization. However, these methods can often lead to image distortion when correcting the saturation component. To address this issue, a method that corrects the saturation component for enhancement is employed:

$$S_c = S + t(V_c - V)\varepsilon \tag{14}$$

In the equation, S_c and V_c represent the enhanced saturation component and luminance component, respectively. t is a constant, and ε is an adjustment coefficient, which can be expressed as:

$$\varepsilon(x, y) = \frac{\sum_{i,j\in\Omega} \left(\left|(V(i,j) - \overline{V_\Omega}(i,j))\right| \left|(S(i,j) - \overline{S_\Omega}(i,j))\right|\right)}{\sqrt{\delta_V(x,y)\delta_S(x,y)}} \tag{15}$$

In the equation, (x, y) represents the enhanced point position. $\overline{V_\Omega}(i,j)$ and $\overline{S_\Omega}(i,j)$ represent the average luminance and saturation values of all points within an N*N neighborhood centered around the enhanced point. $\delta_V(x,y)$ represents the luminance variance of the enhanced point, and $\delta_S(x,y)$ represents the saturation variance of the enhanced point. (i,j) represents the coordinates of the pixels within the neighborhood.

V. EXPERIMENTAL SETUP AND ANALYSIS

To validate the effectiveness of the proposed algorithm for enhancing images in underground mines, quantitative and qualitative comparisons were conducted between the proposed HSV-based enhancement network and existing state-of-the-art algorithms. Additionally, extensive ablation experiments were performed on the network model to demonstrate the improvements brought by the residual multi-scale module and CBAM module, as well as the effectiveness of the designed random luminance perturbation and loss function.

A. EXPERIMENT DETAILS

1) DATASETS

The low-ligh image data in this study was collected from the monitoring footage of the Xieqiao Mine Intelligent Mine System in Huainan City. The dataset consists of a total of 1100 pairs of low/normal light images, which were resized to 512*512 pixels and named the MINE dataset. It is worth noting that we selected only 900 low-light images as the training set and 200 images as the test set. The trained network was also evaluated on seven other datasets, including LOL [42](500 pairs of low/normal light images), LSRW [43] (6650 pairs of low/normal light images, with 50 test images), NPE [44] (8 low-light images), MEF [45] (17 low-light images), LIME [46] (10 low-light images), DICM [47] (69 low-light images), and VV (24 low-light images). LSRW is a large-scale real-world paired low/normal light image dataset, while the latter five datasets only contain low-light images without paired normal light images for reference.

2) IMPLEMENTATION DETAILS

The algorithm in this study was implemented using the PyTorch framework on a PC equipped with both a CPU (Intel(R) Core(TM) i7-12700K) and a GPU (2080TI). The network model was constructed with a batch size of 8. The weight parameters of each layer's filters were initialized

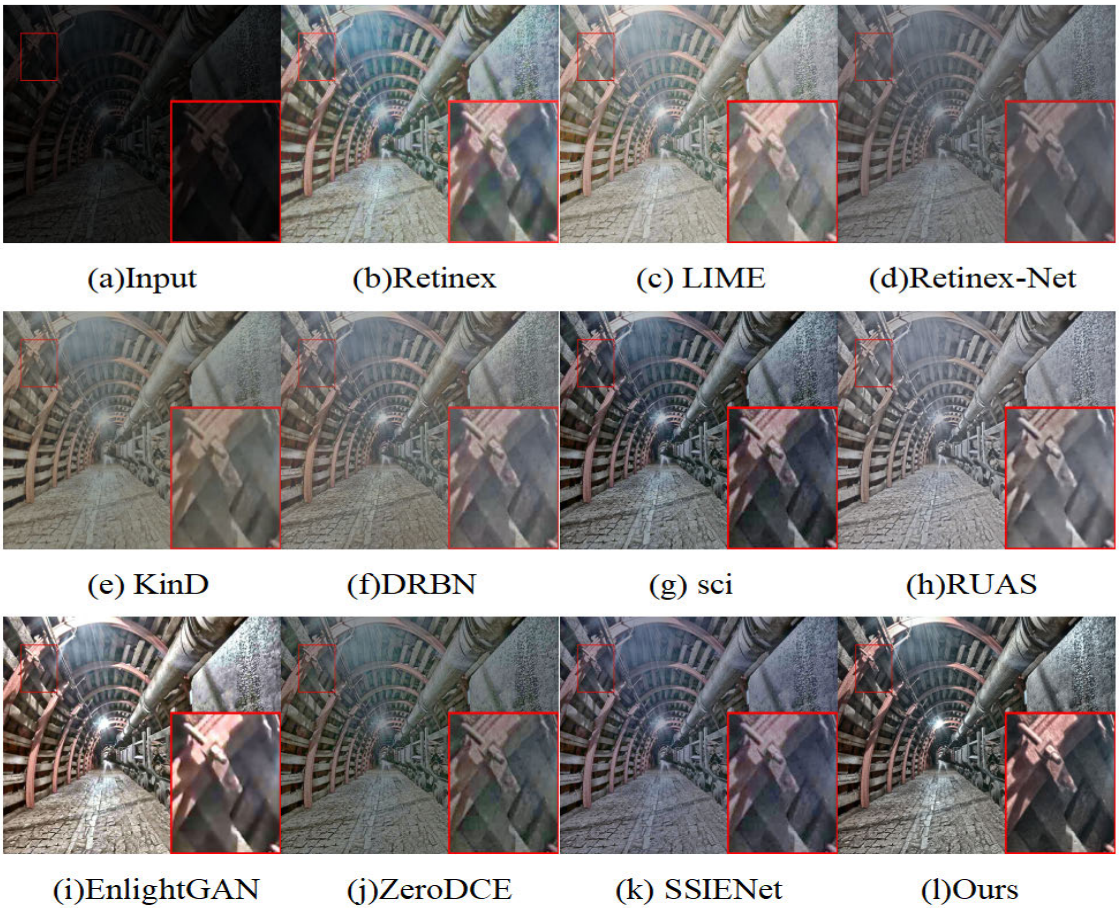

FIGURE 8. Visual comparison among state-of-the-art low-light image enhancement approaches on a low-light image from the mine dataset. Please zoom in for details.

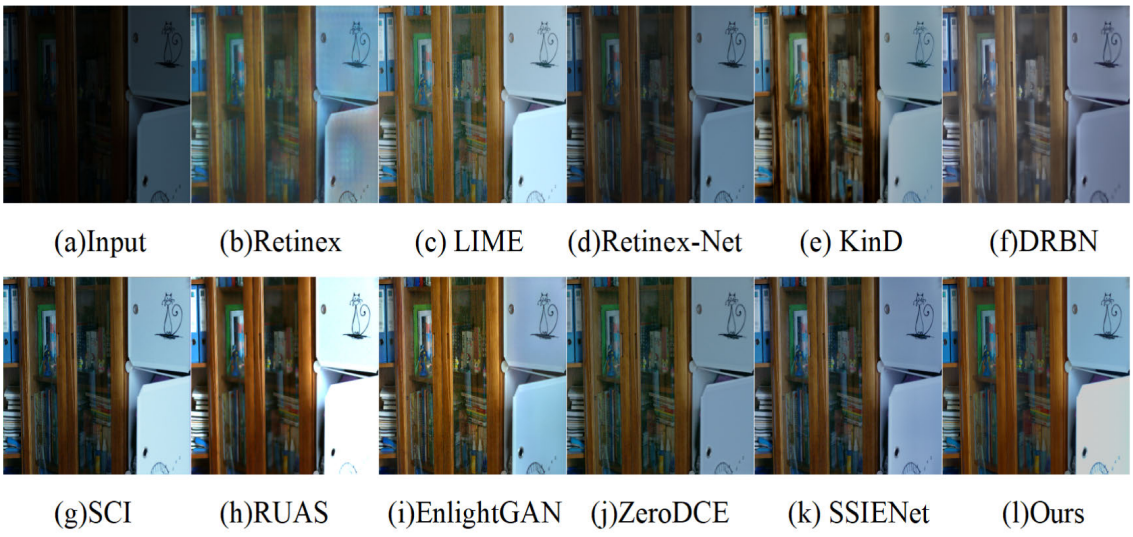

FIGURE 9. Visual comparison among state-of-the-art low-light image enhancement approaches on a low-light image from the LOL dataset. Please zoom in for details.

using the Xavier initialization [48]. The biases were initialized as constants. We employed the ADAM optimizer with default parameters and set the learning rate to 10^{-4}. To introduce brightness perturbations, we applied a weight

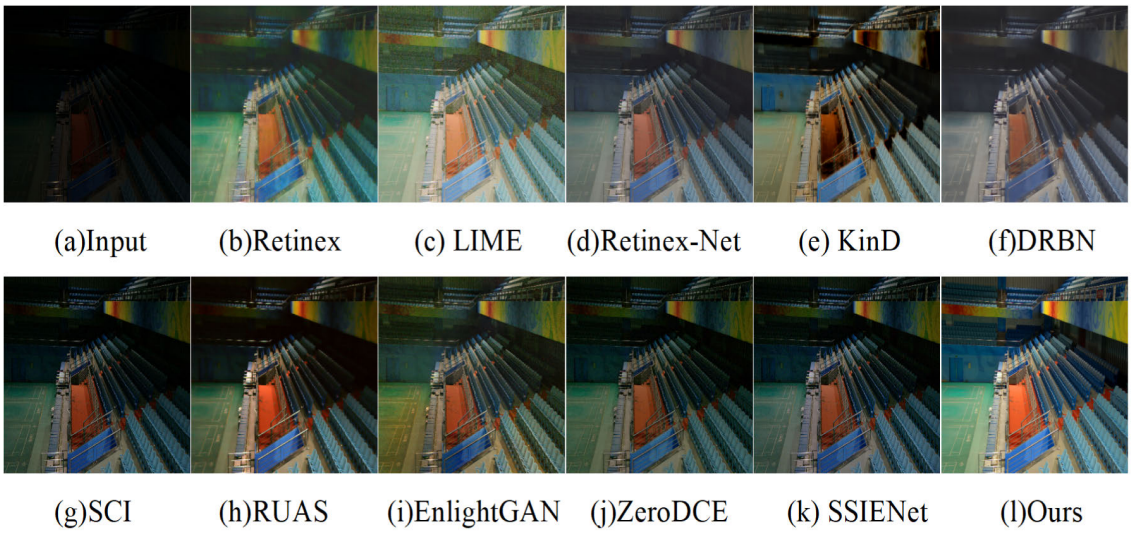

FIGURE 10. Visual comparison among state-of-the-art low-light image enhancement approaches on a low-light image from the LOL dataset. Please zoom in for details.

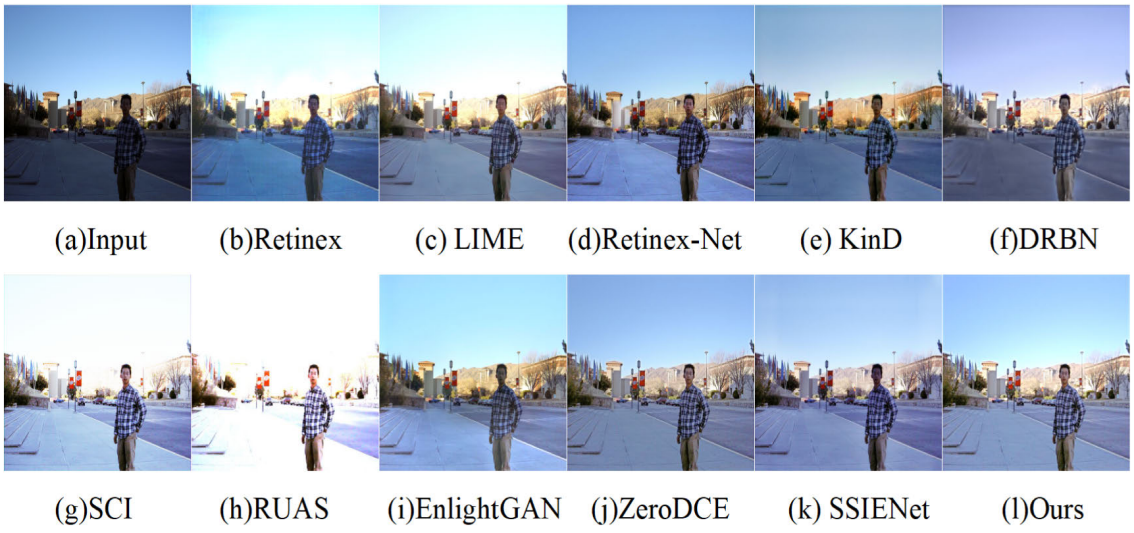

FIGURE 11. Visual comparison among state-of-the-art low-light image enhancement approaches on a low-light image from the DICM dataset. Please zoom in for details.

parameter w of 10. Our method was trained for 100 epochs, and evaluations were conducted every 5 epochs. The best model was selected as the final model for further analysis.

3) EVALUATION METRICS

In this study, two classic full-reference image quality metrics, Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) [49], were used to quantitatively evaluate the enhancement results for datasets that had reference images available, such as MINE, LOL, and LSRW. On the other hand, for datasets that only consisted of low-light images without reference images, two non-reference image quality metrics were employed, including Natural Image Quality Evaluator (NIQE) [50] and Lightness Order Error (LOE) [44]. Higher values of SSIM and PSNR indicate better image quality, while for NIQE and LOE, higher values indicate lower image quality.

4) COMPARED METHODS

The proposed method has been compared with five unsupervised learning-based methods, two traditional methods, and three typical supervised learning methods. The five unsupervised learning-based methods are as follows: 1. EnlightGAN [37]: This method is based on an unsupervised generative adversarial network with a global-local discriminator structure. 2. RUAS [51]: It is built upon the Retinex rule and discovers a low-light prior architecture from a compact search space. SCI [52]: It utilizes a self-calibrating

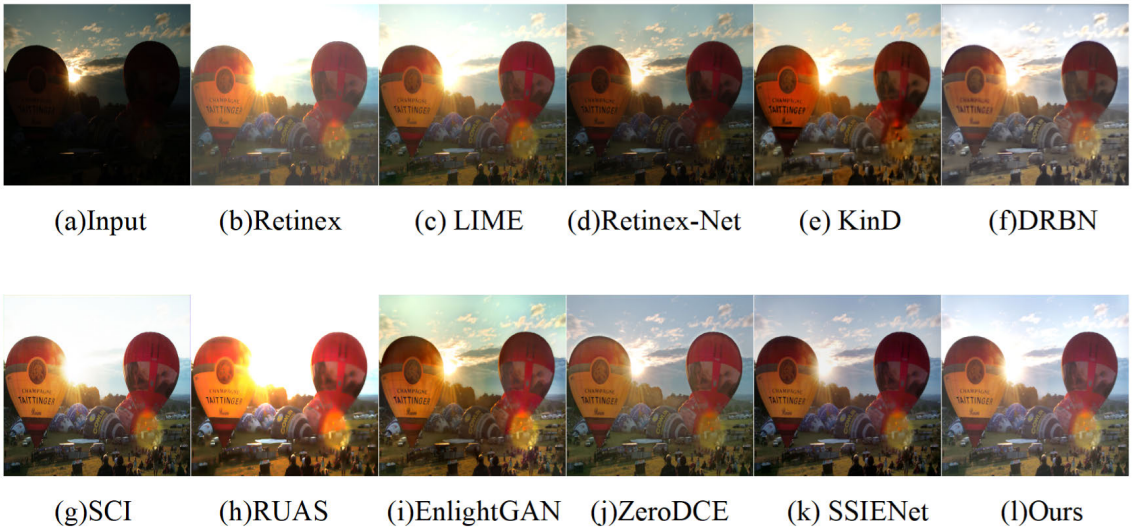

FIGURE 12. Visual comparison among state-of-the-art low-light image enhancement approaches on a low-light image from the VV dataset. Please zoom in for details.

illumination learning framework. 3. ZeroDCE [38]: This method dynamically adjusts the brightness of images by training a high-level curve. SSIENet [53]: It introduces a constraint based on maximum entropy to enhance the performance. The two traditional methods are: 1. Retinex [11]: This method is based on the Retinex theory, aiming to separate the illumination and reflectance components of an image. 2. LIME [46]: It stands for "Local Interpretable Model-agnostic Explanations" and provides explanations for predictions of any machine learning model. The three typical supervised learning methods include: 1. KinD [16]: This method focuses on low-light image enhancement using a deep learning framework. 2. Retinex-Net [25]: It employs a deep neural network to enhance low-light images based on the Retinex theory. 3. DRBN [54]: This method utilizes a deep residual bidirectional network for image enhancement tasks.

B. QUALITATIVE EVALUATIONS

To compare the visual performance of different low-light image enhancement methods, we conducted an extensive evaluation. Figure 8 shows the enhanced results using different methods on the MINE dataset. From Figure 8(a), it can be observed that the original image is overall dark, with low illumination, and unclear image details. Among the traditional algorithms, the Retinex and LIME methods enhance the overall brightness of the image, but the increased brightness leads to overexposure of the overall image. Among the supervised learning methods, the Retinex-Net method lacks sufficient brightness improvement and has low contrast. The KinD and DRBN methods result in blurred texture, as seen in the red-boxed nut contour in Figure 8(e). Among the unsupervised learning methods, the SCI and SSIENet methods still produce dark-colored images after enhancement. The RUAS, EnlightGAN, and ZeroDCE methods enhance the overall brightness and contrast of the images, but the colors appear unrealistic, and the details and contours become blurry. Our method, on the other hand, avoids both underexposure and overexposure issues and effectively restores details and colors, as evident in the red-boxed nut contour in Figure 8(I), which is clearly visible.

Figures 9 and 10 provide visual comparisons of the enhancement results for two real low-light images from the LOL dataset. The experimental results demonstrate that our proposed method exhibits the most natural and realistic color tones, indicating that our model can better restore the image colors. It is particularly noteworthy that the traditional algorithms exhibit varying degrees of overexposure. Other methods either lack sufficient brightness improvements, such as SCI and SSIENet, or introduce color distortions, such as KinD and RUAS. Some algorithms also suffer from blurriness in the details.

Figures 11 and 12 provide visual comparisons of the enhancement results for two real low-light images from the DICM and VV datasets. These comparisons clearly show that previous methods seem to exhibit incorrect exposure, color distortion, noise amplification, or artifacts, which degrade the overall visual quality. For example, the RUAS method produces severe overexposure, ZeroDCE, and SSIENet generate artifacts and amplify noise, while EnlightGAN exhibits color distortion. In contrast, our method consistently produces visually pleasing results, improving both color and brightness without overexposure or underexposure.

C. QUANTITATIVE EVALUATION

Table 1 reports the comparison between our network and other state-of-the-art methods on the MINE, LOL, and LSRW datasets. For each low-light image to be enhanced, there is a corresponding normal-light image available in the

TABLE 1. Quantitative comparisons of different methods on three benchmark datasets.The optimal results are highlighted in red.

Dataset	Metrics	Conventional Methods		Supervised Methods			Unsupervised Methods					
		Retinex	LIME	Retinex -Net	KinD	DRBN	SCI	RUAS	Enlight GAN	Zero DCE	SSIENet	Ours
MINE	PSNR↑	13.32	15.38	16.56	17.62	17.34	16.55	16.11	18.46	16.64	19.15	20.49
	SSIM↑	0.43	0.48	0.53	0.53	0.57	0.53	0.54	0.67	0.57	0.68	0.74
LOL	PSNR↑	14.31	15.76	16.21	17.38	17.55	17.44	16.37	17.65	16.11	19.37	20.31
	SSIM↑	0.46	0.57	0.51	0.55	0.61	0.62	0.52	0.59	0.54	0.71	0.72
LSRW	PSNR↑	13.98	15.83	15.94	16.54	16.53	16.27	15.97	18.58	16.23	17.65	18.15
	SSIM↑	0.44	0.52	0.48	0.51	0.53	0.57	0.53	0.62	0.49	0.53	0.65

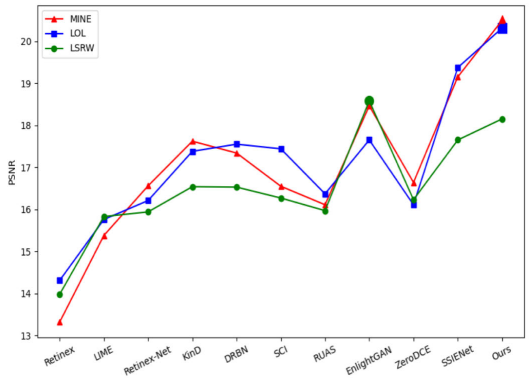

FIGURE 13. PSNR values of the data test image enhancement results.

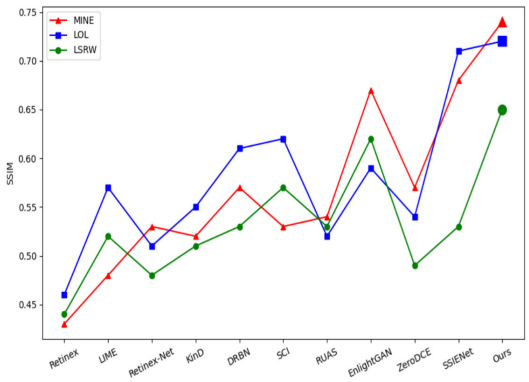

FIGURE 14. SSIM values of the data test image enhancement results.

aforementioned three datasets, allowing us to calculate the PSNR and SSIM values of the enhanced results concerning the normal-light reference images. From Table 1, it can be observed that our proposed model achieves the highest PSNR and SSIM values among all the unsupervised learning methods on the MINE and LOL datasets.

From Figures 13 and 14, it is easier to see that our method achieves the highest SSIM measure on all four datasets, indicating that our method has better structural restoration capability. Additionally, in the LSRW real dataset, the proposed method's PSNR value ranks second, which may be because the images in the LSRW real dataset are not as dark, resulting in a slightly overexposed appearance in the generated images by our method. In fact, our training data mainly consists of darker images from underground mines. To achieve better results, we can consider incorporating a more diverse range of low-light images in the training process. However, overall, the results indicate that our algorithm can maintain high structural similarity and image quality in image enhancement tasks.

Table 2 presents the quantitative results of two no-reference quality assessment metrics, NIQE and LOE, on the NPE, MEF, LIME, DICM, and VV datasets. As shown in Table 2, our proposed method demonstrates competitive performance. Furthermore, although the objective numerical metrics of our method might occasionally be lower than those of other methods, its visual performance is actually superior. Visual comparisons and quantitative analyses of NIQE and LOE can be found in Figure 15. Our method effectively preserves image details and exhibits distinct visual features.

D. REAL-WORLD APPLICATIONS

To evaluate the practicality of the algorithms, real-time images were collected from the monitoring system of the Xieqiao Mine in Huainan City for a comparative enhancement experiment. The captured image is shown in Figure 16(a), where the overall environment is relatively dark, and there is low contrast between the personnel and the background. The experimental results are depicted in Figure 16, where severe overexposure is observed in the Retinex, LIME, and RUAS methods. The Retinex-Net and KinD methods show minimal improvement in brightness after enhancement. The DRBN, SCI, EnlightGAN, ZeroDCE, and SSIENet methods exhibit varying degrees of detail blurring (as indicated by the red boxed area in the image) and color shifts, resulting in a decrease in realism and naturalness of the images. In contrast, our method significantly enhances the overall brightness level of the image, with noticeable brightness improvements in specific regions. The enhanced images exhibit no halo or shadow artifacts, maintain natural and undistorted colors, preserve image information, and exhibit clear details. Thus, our method effectively achieves low-light image enhancement in underground conditions.

E. ABLATION STUDIES

To validate the effectiveness of the proposed improvement method, this study conducted ablation experiments to

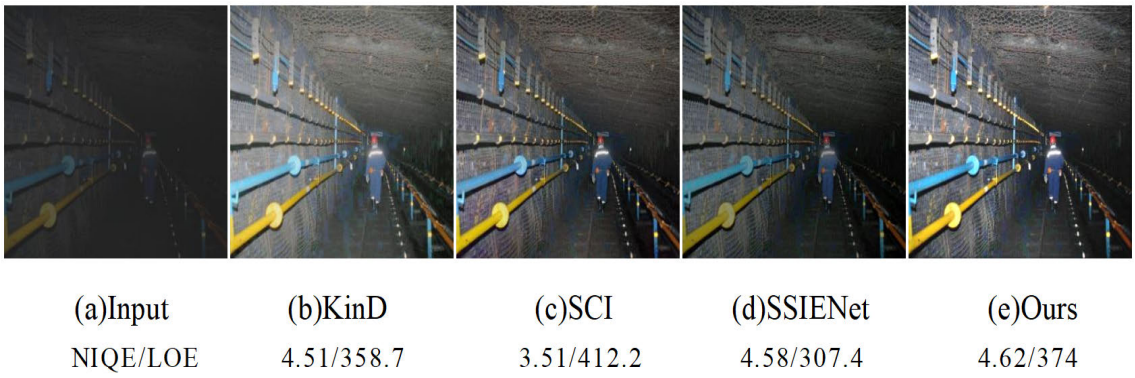

FIGURE 15. Visual comparison of a real-world image among state-of-the-art low-light image enhancement approaches. Please zoom in for details. Compared to other methods, though the objective metrics of the proposed method are inferior, the visual performance is better.

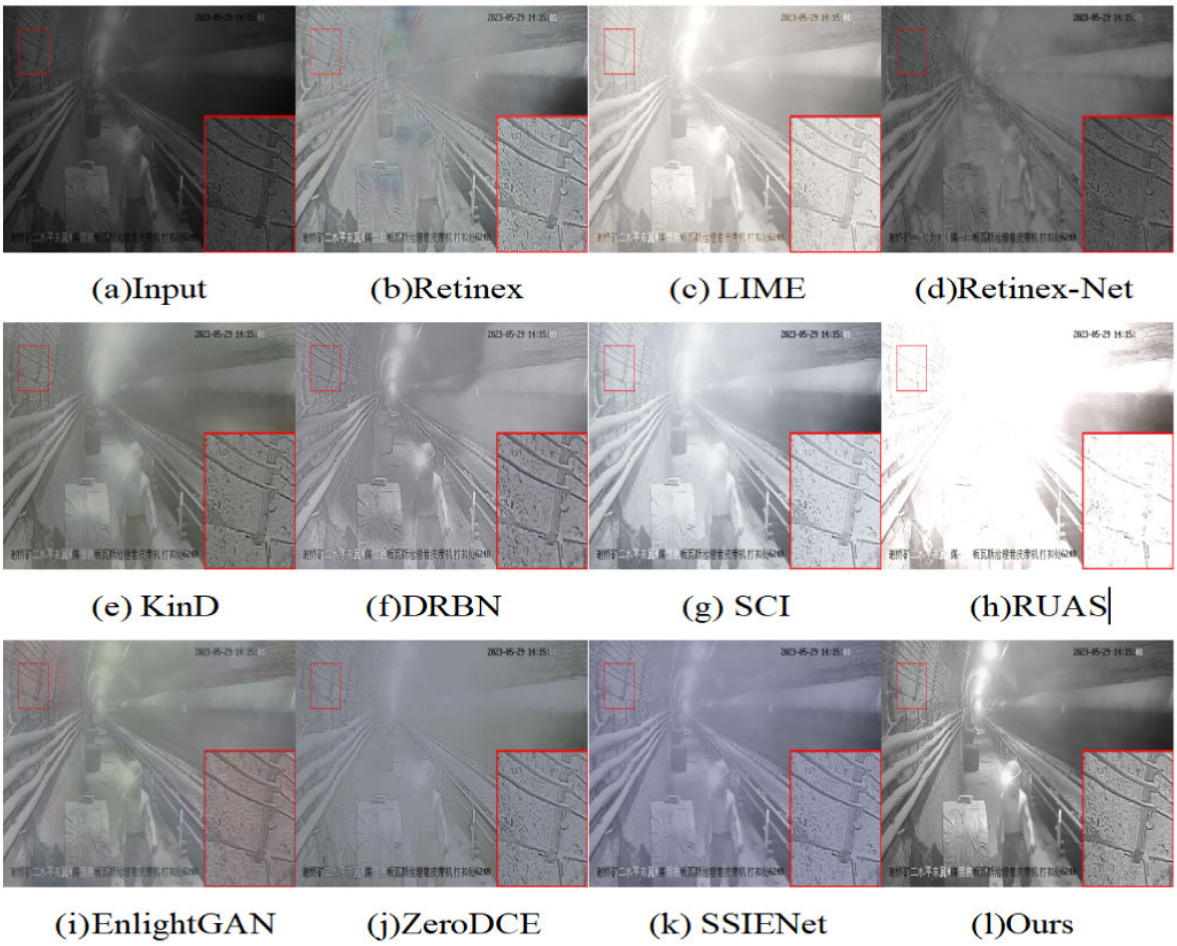

FIGURE 16. Visual comparison among state-of-the-art low-light image enhancement approaches on a low-light image from monitoring system of the Xieqiao Mine in Huainan city. Please zoom in for details.

compare and analyze the results. The experimental results are presented in Figure 17, where (a) represents the low-light input images, (b) shows the test results obtained using our algorithm, (c) displays the test results obtained by removing the consistency of reflectance, (d) shows the test results obtained by removing exposure loss, (e) illustrates the test results obtained by removing structural consistency loss, (f) presents the test results obtained by removing illumination smoothness loss, (g) demonstrates the test results obtained by removing the spatial-channel attention mechanism CBAM,

TABLE 2. Quantitative comparisons of different methods on five benchmark datasets.The optimal results are highlighted in red.

Dataset	Metrics	Conventional Methods		Supervised Methods			Unsupervised Methods					
		Retinex	LIME	Retinex -Net	KinD	DRBN	SCI	RUAS	Enlight GAN	Zero DCE	SSIENet	Ours
NPE	NIQE↓	4.06	4.15	4.68	4.58	4.36	3.99	4.88	3.94	4.02	4.52	3.75
	LOE↓	535.1	872.5	593.4	346.4	439.4	308.8	554.8	398.6	304.9	287.5	256.7
MEF	NIQE↓	4.78	4.62	4.65	4.01	4.58	4.52	4.35	3.88	4.23	4.48	3.98
	LOE↓	543.5	698.6	778.3	423.8	545.7	364.8	441.2	374.1	319.4	288.3	287.1
LIME	NIQE↓	4.54	4.13	4.32	4.62	3.94	4.54	4.33	**4.11**	3.98	5.14	4.01
	LOE↓	498.6	774.8	703.4	543.1	687.1	411.2	359.4	482.1	432.1	401.4	332.4
DICM	NIQE↓	4.68	4.87	4.57	4.03	4.12	4.33	4.12	4.22	4.32	5.05	3.87
	LOE↓	675.8	682.1	554.7	397.4	623.1	397.4	345.6	345.8	384.1	324.3	358.6
VV	NIQE↓	4.14	3.21	3.11	3.15	3.13	2.98	3.17	2.69	3.07	3.55	3.12
	LOE↓	485.6	357.9	398.6	436.1	412.3	291.4	331.2	384.5	409.4	301.1	224.9

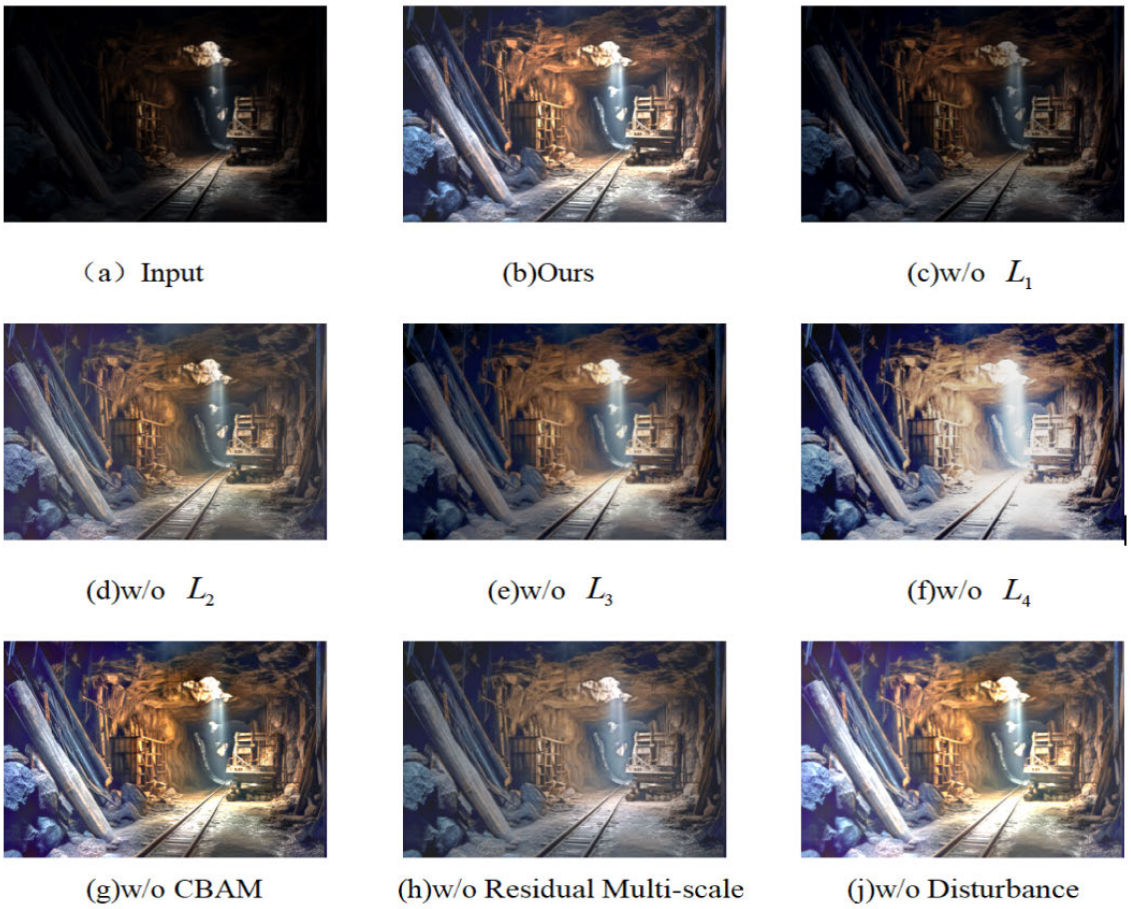

FIGURE 17. Visual comparison among state-of-the-art low-light image enhancement approaches on a low-light image from monitoring system of the Xieqiao mine in Huainan city. Please zoom in for details.

(h) exhibits the test results obtained by removing the multi-scale residual blocks, and (j) showcases the test results obtained by removing random perturbation.

For the (c), (d), (e), and (f) groups, the experiment maintained all modules in the network unchanged while conducting ablation experiments on the loss functions. For the (h), (j), and (k) groups, the experiment kept the loss functions unchanged while conducting ablation experiments on various modules in the network. The test images used for the selected evaluation metrics were obtained from the MINE dataset, which consists of low-light images captured in underground mines. The comparative experimental results are depicted in the figure, and the objective metric evaluations are summarized in Table 3.

From Figure 17(c), it can be observed that in the experiment results where the reflectance consistency loss is

TABLE 3. The ablation experiments correspond to the PSNR and SSIM metrics.

Method	PSNR	SSIM
Ours	**20.39**	**0.672**
w/o L_1	13.48	0.379
w/o L_2	18.26	0.653
w/o L_3	17.39	0.652
w/o L_4	15.77	0.562
w/o CBAM	20.14	0.669
w/o Residual Multi-scale	18.42	0.674
w/o Disturbance	18.41	0.642

removed, the overall color of the image becomes yellowish, and there is abnormal overexposure in the upper-middle region of the image. Despite the introduction of interference, the absence of reflectance consistency loss leads to the network essentially mapping the image-to-image instead of image-to-reflectance. Figure 17(d) shows that after removing the exposure control loss, the image appears globally darker. In Figure 17(e), when the structural loss is removed, the image exhibits some artifacts and loses certain texture details. Figure 17(f) reveals color distortion and severe overexposure near the white light when the illumination smoothness loss is removed. However, Figure 17(g) demonstrates that removing the CBAM module results in exposure issues and partial blurring of details, as seen in the middle of the road in the image. This indicates that the CBAM module allows the network model to focus on important detailed features while suppressing redundant features. Figure 17(h) shows that the absence of the multi-scale residual module leads to a decrease in the network model's feature extraction capability, resulting in an overall whitish appearance of the image. From Figure 17(g), it can be observed that removing random perturbation leads to excessive enhancement, causing halo artifacts and significant distortion at the image edges.

According to the objective evaluation metrics in Table 3, it is evident that regardless of which module is removed, the PSNR and SSIM metrics of the image decrease. It is only when all these modules are combined that the optimal subjective visual effect and objective evaluation results are achieved, thereby validating the effectiveness of each module.

VI. CONCLUSION

The article proposes a coal mine image enhancement method based on the HSV color space, focusing on addressing the impact of dust and low lighting conditions on the enhancement of video images in coal mines. This method combines the characteristics of the HSV color space and applies the Retinex theory to the brightness (V channel). Specifically, the article designs a U-Net-based reflectance estimation network, which improves the algorithm's generalization ability and accuracy by enforcing consistency between the original reflectance and randomly perturbed reflectance. To prevent overfitting, the article designs corresponding loss functions and regularizes the brightness. To further improve the algorithm's performance, residual multi-scale and attention mechanism modules are introduced, which not only enhance the accuracy of the algorithm but also reduce the number of parameters in the network model. During the image enhancement process, the article also adaptively adjusts the saturation (S channel) based on the correlation coefficient to address issues such as reduced saturation and blurred boundaries in the enhanced well wall images. Extensive experiments demonstrate that the proposed method achieves good image enhancement results and provides an effective solution to the image quality issues caused by dust and low lighting conditions in coal mines.

The article mentions the use of the U-Net architecture for reflectance estimation and the incorporation of residual multi-scale and attention mechanism modules. However, regarding the choice of network structure, there are still other possible options and improvement methods, such as considering the use of other deep learning architectures or introducing more complex modules. The algorithm was designed and evaluated under specific conditions (low-light mining images), but the diversity of the mining image dataset created may be limited. If the dataset has a limited number of samples or does not cover various lighting conditions and image features, the training and evaluation results of the algorithm may not be optimal, as mentioned earlier, our proposed method achieved the second-highest PSNR value on the LSRW real dataset.

In conclusion, this article proposes an unsupervised image enhancement method in the HSV space and achieves certain results in the application of low-light mining images. However, in future work, the following aspects should be considered:

1)Multi-modal image enhancement: In addition to low-light mining images, there may be other types of image data, such as infrared images or multi-spectral images. Combining multi-modal image enhancement with low-light mining image enhancement can provide more comprehensive and accurate image enhancement effects.

2)Data augmentation and synthesis: Obtaining real low-light mining images can be challenging and costly. Therefore, data augmentation techniques or data synthesis methods can be considered to generate more training samples. This can improve the model's generalization ability and robustness.

3)Joint optimization and task correlation: Low-light mining image enhancement may be closely related to other tasks, such as object detection or image segmentation. Future research can explore methods for joint optimization, combining image enhancement with other tasks to obtain more comprehensive solutions.

REFERENCES

[1] G. F. Wang, "2025 scenarios and development path of intelligent coal mine," *J. China Coal Soc.*, vol. 43, no. 2, pp. 295–305, 2025.
[2] D. Enjie, Y. Xiao, and X. Bing, "Development of mine informatization and key technologies of intelligent mines," *J. China Coal Soc.*, vol. 47, no. 1, pp. 564–578, 2022.

[3] Y. Chu, Y. Huang, X. Zhang, and H. Liu, "SSD image target detection algorithm based on self-attention," *J. Huazhong Univ. Sci. Technol.*, vol. 48, no. 9, pp. 70–75, 2020.

[4] F. Liu, S. Yang, and H. Yang, "Multi-scale fusion of background and objectness priors for salient object detection," *J. Huazhong Univ. Sci. Technol.*, vol. 46, no. 3, pp. 48–51, 2018.

[5] S. Li, G. Xue, X. Fang, Q. Yang, C. He, S. Han, and Y. Kang, "Coal mine intelligent safety system and key technologies," *J. China Coal Soc.*, vol. 45, no. 6, pp. 2320–2330, 2020.

[6] L. Yan and M. H. Dong, "Multi-layer convolutional features with channel attention for object tracking," *J. Huazhong Univ. Sci. Technol.*, vol. 9, no. 47, pp. 90–94, 2019.

[7] M. Wang and Z. Tian, "Mine image enhancement algorithm based on nonsubsampled contourlet transform," *China Coal Soc.*, vol. 45, pp. 3351–3362, Jan. 2020.

[8] Y. Xu, J. Wen, L. Fei, and Z. Zhang, "Review of video and image defogging algorithms and related studies on image restoration and enhancement," *IEEE Access*, vol. 4, pp. 165–188, 2016.

[9] J. B. Zimmerman, S. M. Pizer, E. V. Staab, J. R. Perry, W. McCartney, and B. C. Brenton, "An evaluation of the effectiveness of adaptive histogram equalization for contrast enhancement," *IEEE Trans. Med. Imag.*, vol. MI-7, no. 4, pp. 304–312, Dec. 1988.

[10] S. Lee, S. Yun, J.-H. Nam, C. S. Won, and S.-W. Jung, "A review on dark channel prior based image dehazing algorithms," *EURASIP J. Image Video Process.*, vol. 2016, no. 1, pp. 1–23, Dec. 2016.

[11] E. H. Land and J. J. McCann, "Lightness and Retinex theory," *J. Opt. Soc. Amer.*, vol. 61, no. 1, pp. 1–11, 1971.

[12] Z. Tian, W. Man-Li, and Z. Yuan-Gang, "Image enhancement algorithm based on dual domain decomposition," *Acta Electonica Sinica*, vol. 48, p. 1311, Jan. 2020.

[13] L. Liu and H. Wang, "Image enhancement using a nonlinear method with an improved single-scale Retinex algorithm," in *Proc. Int. Conf. Electron., Commun. Control (ICECC)*, Sep. 2011, pp. 2086–2089.

[14] Z. Rahman, D. J. Jobson, and G. A. Woodell, "Multi-scale retinex for color image enhancement," in *Proc. 3rd IEEE Int. Conf. Image Process.*, Sep. 1996, pp. 1003–1006.

[15] X.-F. Jiang, G. Wang, and W.-M. Shen, "A method of color image enhancement using color advanced Retinex," *J. Optoelectron. Laser*, vol. 19, no. 10, pp. 1402–1404, 2008.

[16] Y. Zhang, J. Zhang, and X. Guo, "Kindling the darkness: A practical low-light image enhancer," in *Proc. 27th ACM Int. Conf. Multimedia*, Oct. 2019, pp. 1632–1640.

[17] Y. Zhang, X. Guo, J. Ma, W. Liu, and J. Zhang, "Beyond brightening low-light images," *Int. J. Comput. Vis.*, vol. 129, no. 4, pp. 1013–1037, Apr. 2021.

[18] Y. Wang, R. Wan, W. Yang, H. Li, L.-P. Chau, and A. Kot, "Low-light image enhancement with normalizing flow," in *Proc. AAAI Conf. Artif. Intell.*, 2022, vol. 36, no. 3, pp. 2604–2612.

[19] K. Jiang, Z. Wang, Z. Wang, C. Chen, P. Yi, T. Lu, and C. W. Lin, "Degrade is upgrade: Learning degradation for low-light image enhancement," in *Proc. AAAI Conf. Artif. Intell.*, 2022, vol. 36, no. 1, pp. 1078–1086.

[20] K. Xu, H. Chen, X. Tan, Y. Chen, Y. Jin, Y. Kan, and C. Zhu, "HFMNet: Hierarchical feature mining network for low-light image enhancement," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–14, 2022.

[21] Y. Lu, Y. Guo, R. W. Liu, and W. Ren, "MTRBNet: Multi-branch topology residual block-based network for low-light enhancement," *IEEE Signal Process. Lett.*, vol. 29, pp. 1127–1131, 2022.

[22] Y. Zhuang, Z. Zheng, and C. Lyu, "DPFNet: A dual-branch dilated network with phase-aware Fourier convolution for low-light image enhancement," 2022, *arXiv:2209.07937*.

[23] G.-D. Fan, B. Fan, M. Gan, G.-Y. Chen, and C. L. P. Chen, "Multiscale low-light image enhancement network with illumination constraint," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 11, pp. 7403–7417, Nov. 2022.

[24] S.-H. Gao, M.-M. Cheng, K. Zhao, X.-Y. Zhang, M.-H. Yang, and P. Torr, "Res2Net: A new multi-scale backbone architecture," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 2, pp. 652–662, Feb. 2021.

[25] W. Wu, J. Weng, P. Zhang, X. Wang, W. Yang, and J. Jiang, "URetinex-Net: Retinex-based deep unfolding network for low-light image enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2022, pp. 5901–5910.

[26] K. Lu and L. Zhang, "TBEFN: A two-branch exposure-fusion network for low-light image enhancement," *IEEE Trans. Multimedia*, vol. 23, pp. 4093–4105, 2021.

[27] L.-W. Wang, Z.-S. Liu, W.-C. Siu, and D. P. K. Lun, "Lightening network for low-light image enhancement," *IEEE Trans. Image Process.*, vol. 29, pp. 7984–7996, 2020.

[28] Y. Jin, W. Yang, and R. T. Tan, "Unsupervised night image enhancement: When layer decomposition meets light-effects suppression," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2022, pp. 404–421.

[29] Y. Fu, Y. Hong, L. Chen, and S. You, "LE-GAN: Unsupervised low-light image enhancement network using attention module and identity invariant loss," *Knowl.-Based Syst.*, vol. 240, Mar. 2022, Art. no. 108010.

[30] Z. Ni, W. Yang, H. Wang, S. Wang, L. Ma, and S. Kwong, "Cycle-interactive generative adversarial network for robust unsupervised low-light enhancement," in *Proc. 30th ACM Int. Conf. Multimedia*, Oct. 2022, pp. 1484–1492.

[31] R. Wang, B. Jiang, C. Yang, Q. Li, and B. Zhang, "MAGAN: Unsupervised low-light image enhancement guided by mixed-attention," *Big Data Mining Anal.*, vol. 5, no. 2, pp. 110–119, Jun. 2022.

[32] J. Qiao, X. Wang, J. Chen, and M. Jian, "Low-light image enhancement with an anti-attention block-based generative adversarial network," *Electronics*, vol. 11, no. 10, p. 1627, May 2022.

[33] W. Xu, X. Chen, H. Guo, X. Huang, and W. Liu, "Unsupervised image restoration with quality-task-perception loss," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 9, pp. 5736–5747, Sep. 2022.

[34] Q. Jiang, Y. Mao, R. Cong, W. Ren, C. Huang, and F. Shao, "Unsupervised decomposition and correction network for low-light image enhancement," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 19440–19455, Oct. 2022.

[35] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis.*, Jun. 2017, pp. 2223–2232.

[36] J. Bhattacharya, S. Modi, L. Gregorat, and G. Ramponi, "D2BGAN: A dark to bright image conversion model for quality enhancement and analysis tasks without paired supervision," *IEEE Access*, vol. 10, pp. 57942–57961, 2022.

[37] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, "EnlightenGAN: Deep light enhancement without paired supervision," *IEEE Trans. Image Process.*, vol. 30, pp. 2340–2349, 2021.

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

[39] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, Munich, Germany: Springer, Oct. 2015, pp. 234–241.

[40] C. Li, C. Guo, and C. C. Loy, "Learning to enhance low-light image via zero-reference deep curve estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 8, pp. 4225–4238, Aug. 2022.

[41] R. Wang, Q. Zhang, C.-W. Fu, X. Shen, W.-S. Zheng, and J. Jia, "Underexposed photo enhancement using deep illumination estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 6842–6850.

[42] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep Retinex decomposition for low-light enhancement," 2018, *arXiv:1808.04560*.

[43] J. Hai, Z. Xuan, R. Yang, Y. Hao, F. Zou, F. Lin, and S. Han, "R2RNet: Low-light image enhancement via real-low to real-normal network," *J. Vis. Commun. Image Represent.*, vol. 90, Feb. 2023, Art. no. 103712.

[44] S. Wang, J. Zheng, H.-M. Hu, and B. Li, "Naturalness preserved enhancement algorithm for non-uniform illumination images," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3538–3548, Sep. 2013.

[45] K. Ma, K. Zeng, and Z. Wang, "Perceptual quality assessment for multi-exposure image fusion," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3345–3356, Nov. 2015.

[46] X. Guo, Y. Li, and H. Ling, "LIME: Low-light image enhancement via illumination map estimation," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 982–993, Feb. 2017.

[47] C. Lee, C. Lee, and C.-S. Kim, "Contrast enhancement based on layered difference representation of 2D histograms," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 5372–5384, Dec. 2013.

[48] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1026–1034.
[49] W. Burger and M. J. Burge, *Digital Image Processing*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.
[50] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, May 2012.
[51] W. Yang, W. Wang, H. Huang, S. Wang, and J. Liu, "Sparse gradient regularized deep Retinex network for robust low-light image enhancement," *IEEE Trans. Image Process.*, vol. 30, pp. 2072–2086, 2021.
[52] L. Ma, T. Ma, R. Liu, X. Fan, and Z. Luo, "Toward fast, flexible, and robust low-light image enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 5627–5636.
[53] Y. Zhang, X. Di, B. Zhang, and C. Wang, "Self-supervised image enhancement network: Training with low light images only," 2020, *arXiv:2002.11300*.
[54] W. Yang, S. Wang, Y. Fang, Y. Wang, and J. Liu, "From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 3060–3069.

CHANGLIN WU was born in Fuyang, Anhui, China, in 1993. He received the master's degree in control engineering from Anhui University of Engineering, supervised by Prof. Qigong Chen. He is currently with Huainan Normal University, as an Assistant Professor. His research interests include image enhancement and target detection.

DANDAN WANG was born in Kaifeng, Henan, China, in 1987. She received the Ph.D. degree in navigation, guidance and control from Harbin Engineering University, supervised by Prof. Gannan Yuan. She is primarily affiliated with Huainan Normal University, where she works as a Lecturer. Additionally, she holds a part-time position at the Anyang Institute of Technology. Her research interests include carrier attitude estimation algorithms and nonlinear filtering algorithms.

KAIFENG HUANG was born in 1982. He received the bachelor's degree in automation from Anhui University of Science and Technology (AUST), in 2005, the master's degree in control theory and control engineering from the School of Electrical and Information Engineering, AUST, in June 2008, and the Ph.D. degree in safety technology and engineering from the School of Energy and Safety Engineering, AUST, in 2016. He went to Huainan Normal College, in 2017. In 2017, he joined Huainan Normal College, as a Teacher. In September 2020, he became the Vice Dean (in charge of scientific research) of the College of Mechanical and Electrical Engineering, Huainan Normal College, and the Head of the Department of Railway Transportation. His main research interests include safety monitoring and detection and coal mine electromechanical equipment fault diagnosis.