Low-Light Image Enhancement Algorithm for Underground Coal Mines Based on Retinex Theory

Guangyao Yang[1], Qingqing Ran[2], Lihong Dong[2,*], Naining Wen[2]
[1]SHCCIG Yubei Coal Industry Co., Ltd, Yulin, China
[2]College of Computer Science and Technology, Xi'an University of Science and Technology, Xi'an, China
**Corresponding Author*

Abstract: The underground environment of coal mines is complex, causing issues such as insufficient lighting, low contrast, and high noise in the images collected by underground monitoring equipment. These problems severely affect subsequent image analysis and intelligent decision-making. This paper introduces a low-illumination image enhancement algorithm for underground coal mines, grounded in Retinex theory. Firstly, the projection module is utilized to process the original image, reducing the interference of noise on Retinex decomposition. Secondly, a decomposition network integrated with the U-Net structure is employed to accurately separate the illumination and reflection components. Finally, a self-calibration illumination network is introduced. Through multi-stage residual learning and self-calibration mapping, it can automatically adjust the illumination component. Experiments demonstrate that on the self-built underground image dataset, the proposed algorithm outperforms mainstream methods in terms of PSNR and SSIM metrics. The algorithm presented in this paper outperforms comparative algorithms in subjective human vision analysis, demonstrating its effectiveness in enhancing the visual quality of low-illumination images in underground coal mines.

Keywords: Low Illumination Imaget; Image Enhancement; Retinex Theory; Underground Coalmine.

1. Introduction

In the process of coal mine intelligentization, the underground video surveillance system[1] is of great significance for safe production and efficient operation. It can provide data for tasks such as personnel positioning and behavior recognition, enabling surface personnel to monitor the underground situation in real time and make decisions. However, the underground environment is harsh, with insufficient and unevenly distributed artificial lighting, as well as a large amount of dust and fog. This results in inadequate illumination and detail loss in collected images, diminishing image quality and hindering computer vision tasks like object detection and image recognition, as well as manual observation and intelligent decision-making. Therefore, studying the low-illumination image enhancement technology for coal mines can promote the intelligent development of coal mines, ensure personnel safety, and improve production efficiency, etc.

Recent advancements in image processing technology have led to significant progress in enhancing low-illumination images in coal mines, utilizing both traditional and deep-learning methods. However, many challenges still remain. Traditional image enhancement techniques, such as grayscale transformation-based histogram modification[2], enhance image brightness and contrast by redistributing the pixel histogram. However, this approach often leads to issues like color distortion and local over-enhancement. The Retinex theory-based algorithm[3] enhances images by focusing on the reflection components of objects; however, it frequently produces halo effects and artifacts when applied to images with shadow transitions or artificial light sources. In addition, some methods based on physical models and prior knowledge may be effective under specific conditions, but in complex underground scenarios, they often lose details and suffer from color distortion.

Retinex-Net[4] combines Retinex theory with convolutional neural networks to enhance image brightness via decomposition and enhancement networks, but it faces challenges like color distortion and detail loss. The KinD[5] algorithm, introduced by Zhang et al., is prone to inconsistencies between the reflection and illumination components during network decomposition, leading to the emergence of artifacts. Guo et al.[6] devised the Zero-DCE network, which achieves brightness enhancement through curve iteration. However, it falls short in adequately enhancing brightness for darker, low-light underground images and suffers from overexposure. The Enlighten GAN network, proposed by Jiang et al.[7], employs generative adversarial networks to establish a mapping between unpaired images, yielding results in color restoration and enhancement, yet it still contends with shadowed areas. IAT[8] presents an illumination-adaptive Transformer structure aimed at exposure correction and image enhancement. Nonetheless, its frequent use of convolutional layers and lack of focus on the recovery of edge detail information diminish the quality of its restoration algorithm. SCI[9] presents a self-calibrating illumination framework utilizing a weight-sharing cascaded structure to learn the illumination enhancement process; however, it has limited ability to maintain the detailed features of low-light images. Currently, despite the existence of numerous low-light enhancement algorithms, challenges persist in enhancing images from underground coal mines due to the complex environment, inadequate roadway lighting, limited image acquisition devices, and interference from dust and background. These challenges result in issues such as low contrast, weak edge information, and high image noise. This paper introduces a low-light image enhancement algorithm for underground coal mines, grounded in Retinex theory. The primary contributions include:

1) A projection module is introduced to remove inappropriate noise, ensuring that the decomposition network can stably and accurately decompose the image.

2) A channel-splicing U-Net-based low-illumination image decomposition network is developed to process the original image by extracting and analyzing the interrelationships between the illumination and reflection components at various levels, ultimately decomposing the image into these two components.

3) A self-calibration module is introduced to construct a progressive illumination optimization process, increasing the overall illumination of the image and thus improving the image quality.

2. Low-Light Image Enhancement Algorithm for Underground Coal Mines Based on the Retinex Theory

Low-light images in underground coal mines exhibit greater complexity than typical low-light images. In addition to insufficient brightness, there are interferences from dust and fog, resulting in low contrast and blurred details. Moreover, in the special working environment of underground coal mines, the backgrounds of areas such as the underground working faces are generally dark, which further increases the difficulty of separating the targets from the backgrounds in low-light images. We introduce a low-light image enhancement algorithm for underground coal mines, grounded in Retinex theory. Initially, the image is processed by a projection module to eliminate unwanted noise, facilitating stable and precise image decomposition by the network. Subsequently, the denoised image undergoes decomposition into illumination and reflectance components via a dedicated module. Subsequently, a network is employed to estimate the illumination component's brightness. By precisely analyzing brightness, this network offers suitable illumination adjustments to enhance the image's visual quality. The illumination map is integrated with the reflectance map using Retinex theory to improve contrast and detail in underground mine images. Figure 1. shows the overall network structure of this paper.

2.1 Denoising Network

In the low-light image enhancement task, actual low-light images are not noise-free images in an ideal state. The accuracy of Retinex decomposition is vulnerable to interference from non-ideal features such as noise in the original image, local overexposure, or high-frequency disturbances. To address this issue, inspired by Fu et al.[10], this paper introduces the projection module P-Net. Its core objective is to generate a projected image suitable for Retinex decomposition through feature remapping. As depicted in Figure 2., P-Net is composed of four 3x3 convolutional layers with ReLU activation, followed by an additional 3x3 convolutional layer. The last layer normalizes the output to the interval [0, 1] through the Sigmoid function, and the number of output channels is 3 to ensure alignment with the color space of the original image. The role of P-Net can be explained from the perspective of error redistribution. The objective function optimized during image decomposition by the network is expressed as:

$$\begin{aligned} &\underset{L,R}{\arg\min} \| Y - LR - \varepsilon \| \\ &= \underset{L,R}{\arg\min} \| Y - T + T - LR - \varepsilon \| \\ &\leq \| Y - T + T \| + \underset{L,R}{\arg\min} \| T - LR \| \end{aligned} \tag{1}$$

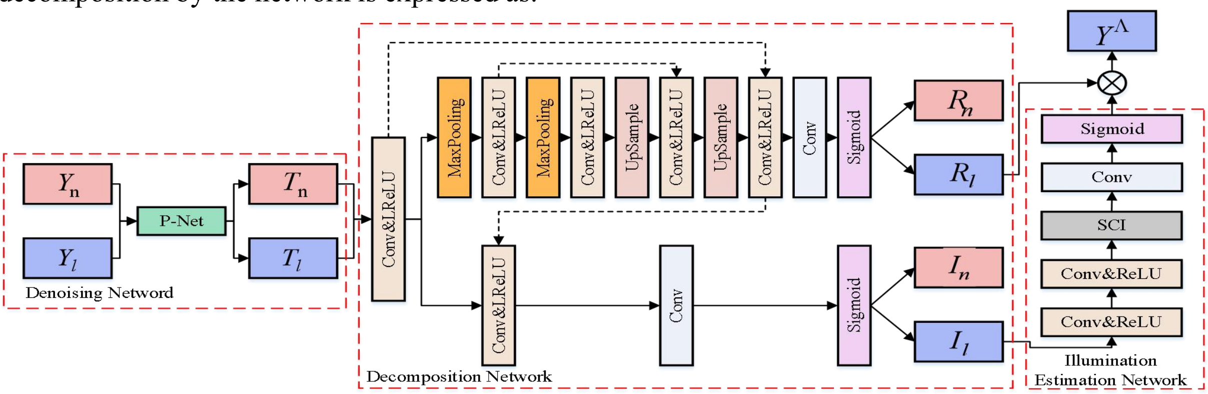

Figure 1. Overall Network Structure

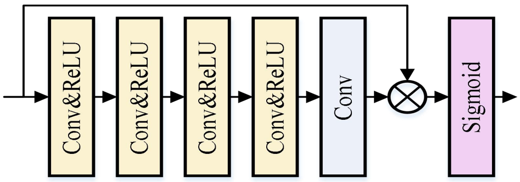

Figure 2. P-Net Network Structure

Among them, represents the original low-light image, represents the projected image, represents the error in the original image caused by noise factors, and and represent the decomposed illumination component and reflectance component, respectively. It can be seen from (1) that the decomposition process will be affected by errors that cannot be modeled by the Retinex theory. The P-Net performs projection preprocessing on the original image, transferring the errors from the decomposition stage to the projection stage. The method enhances the Retinex decomposition task by preserving the similarity between the projected and original low-light images while effectively eliminating unwanted noise.

The projection module's loss function is formulated as:

$$L_P = \|Y - T\|_2 \tag{2}$$

2.2 Decomposition Network

The decomposition network module consists of two branches dedicated to extracting the reflectance and illumination maps, respectively. The reflectance map branch adopts the U-Net structure. Two 3×3 convolutional layers and the Leaky ReLU function are used to capture the contextual features. A 2×2 max pooling layer follows the convolutional block. The upsampling process corresponds to the downsampling process. Skip connections are used to perform feature fusion on the feature maps at different levels. In this way, not only the high-resolution detail information contained in the high-level feature maps is retained, but also the problem of the loss of shallow features caused by the increase in the depth of the model is avoided, thus achieving a high-quality image output. The structure used to extract the illumination map is formed by concatenating with the upsampling and downsampling structures of the upper branch, followed by a Conv+Leaky ReLU layer and a Conv layer, and finally a Sigmoid layer. This concatenation provides additional guiding information for the lower branch to estimate the illumination component more accurately. Figure 1 illustrates the overall structure.

The decomposition module's loss function is formulated as:

$$L_{\mathrm{d}} = L_{\mathrm{res}} + \lambda_{\mathrm{r}} L_{\mathrm{ref}} + \lambda_{\mathrm{i}} L_{\mathrm{ill}} + \lambda_{\mathrm{m}} L_{\mathrm{mc}} \tag{3}$$

Where λ_{r}, λ_{i}, and λ_{m} are the weight coefficients.

The reconstruction loss ensures the consistency between the original input image and the product of the decomposed reflectance and illumination components. The reconstruction loss can be expressed as:

$$L_{res} = \|R_l I_l - L_l\|_1 + \|R_n I_n - L_n\|_1 \tag{4}$$

Where R_n , R_l , I_n , I_l are respectively the illumination component and reflectance component after the decomposition of the input image, and $\| \ \|_1$ is the L1-norm.

The reflectance component, indicative of an object's intrinsic properties like texture and color, should remain unaffected by illumination. The reflectance component consistency loss aims to constrain the reflectance maps of the same scene under low-light and normal-light conditions to be consistent. The reflectance component consistency loss can be expressed as:

$$L_{ref} = \|R_l - R_n\|_1 \tag{5}$$

The illumination component smoothness loss ensures that the illumination component is as smooth as possible in texture details while retaining the overall structural information by constraining the spatial gradient change of the illumination component. The formula is:

$$L_{ill} = \left\| \frac{\nabla I_l}{\max(|\nabla I_l|, \varepsilon)} \right\|_1 + \left\| \frac{\nabla I_n}{\max(|\nabla I_n|, \varepsilon)} \right\|_1 \tag{6}$$

Where ∇ is a first order gradient operator, which includes the horizontal and vertical directions. ε is a very small positive constant used to avoid division by zero errors.

Inspired by KinD, the mutual consistency loss is added to ensure the consistency of the illumination component in smooth regions and edge regions. The formula is as follows:

$$L_{mc} = \|M \cdot \exp(-c \cdot M)\|_1 \tag{7}$$

Where, $M = |\nabla L_l| + |\nabla L_n|$.

2.3 Illumination Estimation Network

The illumination adjustment network aims to optimize the brightness distribution of the illumination map. To this end, this paper uses the self-calibrating illumination network (SCI) to process images. The self-calibrating illumination network not only boosts image brightness but also restores texture details effectively. The structure of SCI is shown in Figure 3.

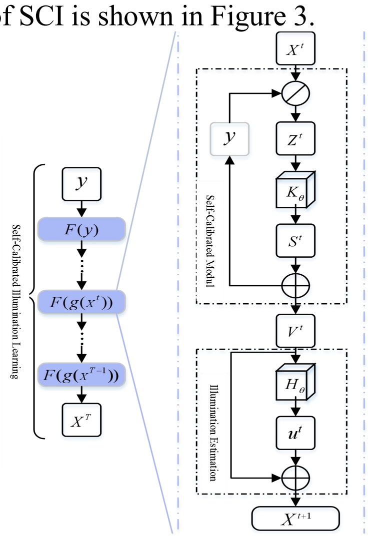

Figure 3. SCI Structure

SCI utilizes staged illumination learning based on Retinex theory to model the relationship between low-light observations and illumination through residual learning. Its basic unit is expressed as:

$$F\left(X^t\right): \begin{cases} u^t = H_\theta\left(X^t\right), X^0 = y \\ X^{t+1} = X^t + u^t \end{cases} \tag{8}$$

Where u^t and X^t represent the residual term and illumination at the t stage; H_θ is the illumination estimation network, which adopts a weight-sharing mechanism to make each stage share the same network structure and parameters. SCI incorporates a self-calibration module into the low-light observations to maintain result consistency across each stage by correcting the input. The self-calibration module is represented by the following formula:

$$G\left(X^t\right): \begin{cases} Z^t = y \otimes X^t \\ S^t = K_\theta\left(Z^t\right) \\ V^t = y + S^t \end{cases} \tag{9}$$

Where Z^t is the fusion result calculated from the output X^t of the current stage and the low-light observation image y, S^t is the self-calibration map, which is generated by the parameterized operator K, and V^t is the calibrated input for the next stage.

Therefore, the basic unit of the t-th stage lighting optimization process can be reformulated as:

$$F\left(X^t\right) \rightarrow F\left(G\left(X^t\right)\right) \tag{10}$$

The algorithm enhances low-illumination image quality and clarity in coal mines by integrating illumination learning with a self-calibration module during the learning process.

3. Experimental Results and Analysis

3.1 Datasets and Experimental Environment

Capturing paired low-light and normal-light images in coal mines is challenging due to the complex environment and limitations of shooting equipment. To address this issue, images are extracted from actual underground monitoring footage and videos shot by intrinsically safe law enforcement recorders in mines through frame extraction. A set of 1000 images with normal illumination was chosen, and low-light images were created by modifying their contrast and brightness to mimic actual low-light conditions in coal mines. The dataset, combined with a portion of the existing low-light dataset (LOL), is utilized to train the network model.

The proposed low-illumination image enhancement algorithm for coal mines, based on

Retinex theory, is implemented in Python to assess its effectiveness in complex coal mine environments. The study utilizes a custom dataset of underground coal mine scenes, with training conducted on an NVIDIA GeForce RTX 3090 GPU. The parameter settings during training are as follows: batch size = 8, epochs = 500, and the input image size is 600×400.

3.2 Subjective Evaluation

The proposed algorithm's effectiveness is assessed by selecting three underground coal mine images from the dataset and conducting a subjective comparative analysis of the experimental results with different algorithms. In this experiment, five methods, namely LIME, Retinex-Net, KinD, Zero-DCE, and IAT, are selected as comparative algorithms.

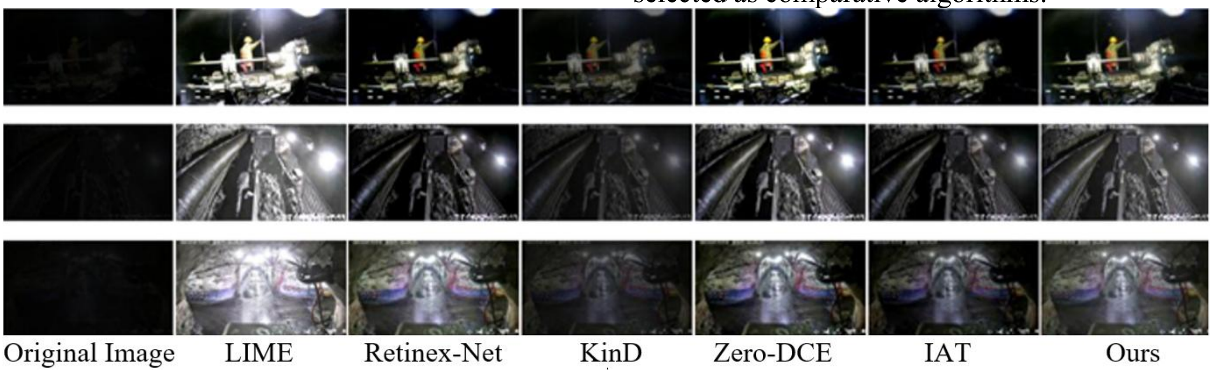

Figure 4. Comparison of Enhancement Results of Different Algorithm

Figure.4. illustrates the enhancement effects of various methods on low-illumination images from the mine. It is not difficult to see from the figure that various methods have optimized the visual performance of the mine images to a certain extent, and the brightness, contrast, and clarity of the images have all been improved. However, after being enhanced by the LIME method, there is an obvious overexposure phenomenon at the point light sources in the images have all been improved. However, after being enhanced by the LIME method, there is an obvious overexposure phenomenon at the point light sources in the image, resulting in an image with enhanced halos. The Retinex-Net method results in color distortion and fails to reveal certain details in dark areas. The KinD method produces images that are somewhat dark, with low contrast and slightly blurred details. The image enhancement effect of the Zero-DCE method is relatively excellent, and the image fidelity is high. However, the illumination distribution in individual areas is still not uniform enough, resulting in an unbalanced color situation. The IAT method will cause a large amount of detail texture of the image to be lost, the image smoothness is too high, and it is difficult to observe the detail information in the image. The algorithm introduced in this paper effectively adjusts the majority of image areas to optimal lighting conditions, surpassing the methods previously discussed. It not only makes the enhanced image have sufficient brightness and avoids over-enhancement, but also enables the information in the dark areas to be displayed normally. The overall brightness and local texture details closely resemble those of an image under normal illumination, resulting in a more ideal visual effect.

3.3 Objective Evaluation

To objectively assess the proposed algorithm's effectiveness, we employ two image quality metrics: Peak Signal-to-Noise Ratio (PSNR) and Structural SIMilarity (SSIM), acknowledging that subjective visual comparisons can differ among individuals. PSNR quantifies image distortion by computing the mean squared error between the output and the reference image. The larger the value, the smaller the distortion of the enhanced image, and the better the image quality. SSIM measures the similarity of images. A higher value indicates less image distortion and greater structural similarity between the enhanced and original images. Refer to Table 1 for detailed results.

Table 1. Objective Evaluation Results of Differ-ent Models

Methods	PSNR	SSIM
LIME[3]	16.37	0.532
Retinex-Net[4]	19.62	0.604
KinD[5]	20.49	0.784
Zero-DCE[6]	17.21	0.718
IAT[8]	21.58	0.789
Ours	22.86	0.793

Table 1 demonstrates that the proposed algorithm enhances images, resulting in higher PSNR values compared to other algorithms when evaluating the dataset. An elevated PSNR value signifies that the enhanced image closely resembles the normal illumination image in pixel content. The elevated SSIM parameter value indicates that the algorithm presented in this paper enhances images to more closely resemble normal illumination images in terms of structural characteristics compared to other algorithms. In conclusion, whether it is the evaluation of subjective visual effects or objective indicators, the proposed algorithm has significantly improved in terms of brightness, contrast, clarity, and image detail information. It effectively enhances low-illumination images and offers distinct advantages.

4. Conclusions

The proposed algorithm initially eliminates unsuitable noise using the projection module. Secondly, integrating the decomposition network utilizing the residual U-Net structure enhances the accuracy of separating the illumination and reflection components. Utilizing multi-stage residual learning and self-calibration mapping, the method effectively balances global brightness enhancement with local detail preservation, significantly improving the visual quality of low-light images. Future research will investigate the connection between low-light image enhancement and advanced visual tasks like object detection and image recognition, aiming to further optimize the model and its performance.

References

[1] Deqiang Chen, Jiansheng Qiu, Xingge Guo. Review on key technologies of AI recognition for videos in coal mine Coal Science and Technology, 2023, 51(2):349-365.

[2] A. Paul. Adaptive tri-plateau limit tri-histogram equalization algorithm for digital image enhancement The Visual Computer, 2023, 39(1):297-318.

[3] Xiaojie Guo, Yutao Li, Haibin Ling. LIME: Low-light image enhancement via illumination map estimation IEEE Transactions on Image Processing, 2016, 26(2):982-993.

[4] Yifan Jiang, Xiaowei Gong, Dong Liu, et al. EnlightenGAN: Deep light enhancement without paired supervision IEEE Transactions on Image Processing, 2021, 30: 2340-2349.

[5] Ying Zhang, Jing Zhang, and Xiaojie Guo. Kindling the darkness: A practical low-light image enhancer Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 2019.

[6] Chunle Guo, Chongyi Li, Jinjian Guo, Chen Change Loy, Junhui Hou, and Sam Kwong. Zero-reference deep curve estimation for low-light image enhancement Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2020.

[7] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep Retinex decomposition for low-light enhancement Proceedings of the 2018 British Machine Vision Conference, Newcastle, UK, 2018.

[8] Ziteng Cui, Kunchang Li, Lin Gu, Shenglong Su, Pan Gao, and Zheng Jiang. You only need 90k parameters to adapt light: A lightweight transformer for image enhancement Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, Montreal, Canada, 2022.

[9] Long Ma, Tengyu Ma, Risheng Liu, Xin Fan, and Zhongxuan Luo. Toward fast, flexible, and robust low-light image enhancement Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, USA, 2022.

[10] Zheng Fu, Yang Yang, Xiaogang Tu, Yuhui Huang, Xiaoming Ding, and Kaiming Ma. Learning a simple low-light image enhancer from paired low-light instances Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, Canada, 2023.