Academic Journal of Engineering and Technology Science
ISSN 2616-5767 Vol.4, Issue 2: 52-56, DOI: 10.25236/AJETS.2021.040208

Low-light images enhancement method in coal mine based on knowledge distillation

Caoshuai Kang[1, a,*]

School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China
[a]506954185@qq.com
**Corresponding Author*

Abstract: *In order to solve the problems of low illumination and poor image quality in video monitoring systems caused by complex lighting conditions in underground mines, a low-light image enhancement method based on knowledge distillation is proposed. By establishing Net-S (student network) and Net-T (teacher network) for feature simulation, this method improves the image reconstruction ability of the model while transferring knowledge and improves the network efficiency and image enhancement effect at the same time. The improved RIR (residual in residual blocks) module is added to improve the feature extraction ability of the model. At the same time, the loss function for low illumination image characteristics is designed to improve the accuracy of the network model in image reconstruction. Compared with the traditional knowledge distillation network model, Net-T and Net-S of the model output different images respectively. Net-T extracts features from high-quality images. Net-S simulates features under the guidance of Net-T, and finally obtains enhanced images. The experimental results show that the model can effectively remove part of the noise, improve the brightness and quality of the image, and improve the visual effect of the image. Moreover, the model has good generalization ability. Compared with CLAHE and LIME, the absolute cumulative change of the normalized objective index of this method is increased by 42.21%, 7.47% and 30.18% respectively. In view of the low illumination image enhancement of the mine, this method has certain advantages, which can significantly improve the overall brightness of the image, reduce the noise, and is more in line with the human visual perception. To a certain extent, it meets the needs of the construction of intelligent mine video monitoring system.*

Keywords: *low-light images, image enhancement, knowledge distillation, transfer learning*

1. Introduction

One of the important goals of smart mine construction is to achieve the target that less or no people working in the working face under the mine. The research on intelligent video monitoring technology and its control optimization is of great significance to promote the construction of smart mine and the development of intelligent coal mine safety mining technology [1,2]. The realization of mine intelligent video monitoring technology needs to accurately extract the information in the video image and carry out image segmentation processing. Due to the special low light and dust environment of the mine, some regional monitoring images lose information due to the lack of illumination, which seriously affects the normal visual perception of the staff, and is difficult to be used by the intelligent computer vision system that needs normal illumination image [3]. In order to better present the scene information of coal mine, improve the visual perception of the image, make full use of the existing computer vision system, and promote the development of intelligent video monitoring technology and mine intelligence, the research of low illumination image enhancement method in coal mine is of great significance.

At present, the methods for low illumination image enhancement in coal mine mainly include the following aspects: (1) histogram equalization (HE) and its variants [4]. These kinds of methods output enhanced image by histogram constraints on the output image to meet the constraints. (2) Methods based on Retinex theory [5-6], These kinds of methods assume that the observed image can be decomposed into two aspects of reflectivity and illumination to design parameters and output enhanced images. (3) Deblurring methods [7], These kinds of methods reversely connects the image of insufficient illumination and the image of fogging blur, so as obtaining the enhanced image.

ISSN 2616-5767 Vol.4, Issue 2: 52-56, DOI: 10.25236/AJETS.2021.040208

2. Introduction of knowledge distillation method

Leo Breiman et al. [8] pointed out that different random algorithm programs may lead to different models with similar verification performance, and these models can be combined into a whole to achieve the prediction ability better than each component model. On the contrary, knowledge distillation can distill information from a complex set of models, reaching the performance index of the original model. In deep learning, the neural network is usually used to transform logic information by SOFTMAX layer. In this process, the direct output of neural network is transformed into probability. When the median value of the formula is set to 1, the output layer is hard target distribution, that is, only the type with the highest output probability. The size determines the smoothness of the output probability of the SOFTMAX layer, and the entropy of the distribution is proportional to the information carried by the soft tag. As shown in Figure 1, the temperature changes the degree of attention to soft tags during net-s training.

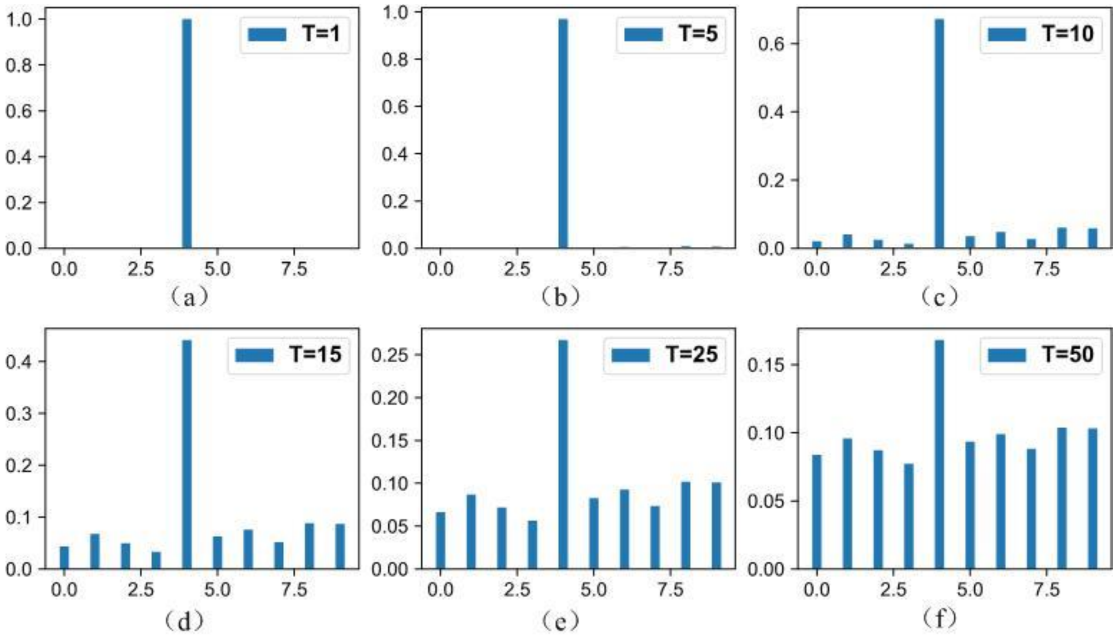

Figure 1: Prediction probability change results of classification prediction experiment under different temperatures.

2.1. Dilated convolution

Dilated convolution is a convolution method proposed by Fisher Yu to improve the accuracy of semantic segmentation model. It can systematically aggregate multi-scale context information of image without losing resolution. The convolution result is shown in Figure 2. Research has proved the effectiveness of hole convolution in image processing. Compared with the traditional convolution method, hole convolution can solve the problems of traditional convolution, such as the determination of up sampling and pooling layer output results, the loss of image internal data structure and spatial information. But from the traditional design idea of hole convolution, hole convolution will lose the continuity of information in the process of multi-scale image information aggregation, ignore the information of small objects, and reduce the feature extraction ability of image information.

The first step is to use the depth convolution of adding dilated rate to convolute the input image features one by one, the second is to combine the convolution layers with different dilated rates to form a mixed convolution module to systematically aggregate the multi-scale context information of the image. Multi scale image feature extraction can solve the mesh effect caused by the cavity convolution and the problem that the feature extraction can't be considered by the objects far and near.

ISSN 2616-5767 Vol.4, Issue 2: 52-56, DOI: 10.25236/AJETS.2021.040208

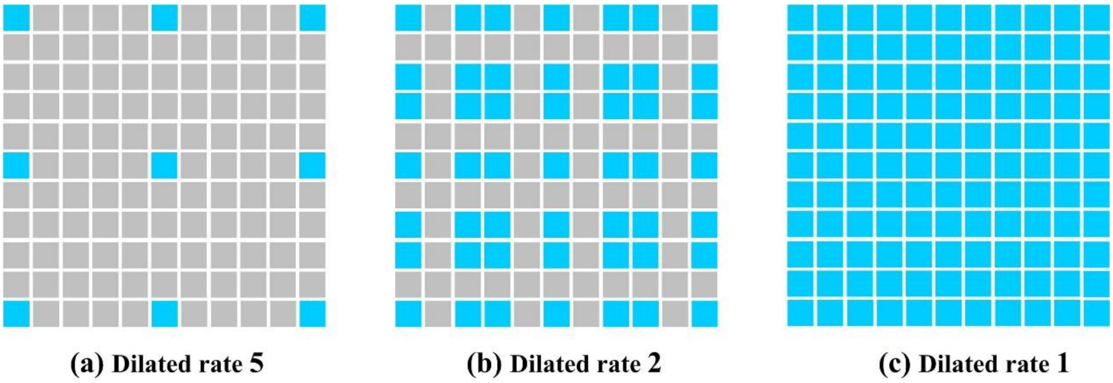

Figure 2: Results of dilated convolution with different dilation rates.

2.1.1. Attention map

Attention map focuses on different image information in different shallow and deep layers. As shown in Figure 3, the neural network pays more attention to the detail texture of the image in shallow layer and pays more attention to the overall information of the image in deep layer.

In the whole structure of mine low illumination image enhancement network model based on knowledge distillation, the network model is divided into net-t and net-s (image enhancement network). Net-t adopts the automatic encoding and decoding structure in [9], and net-s adopts the network structure similar to net-t. The net-s coding part adopts partial hybrid dilated convolution.

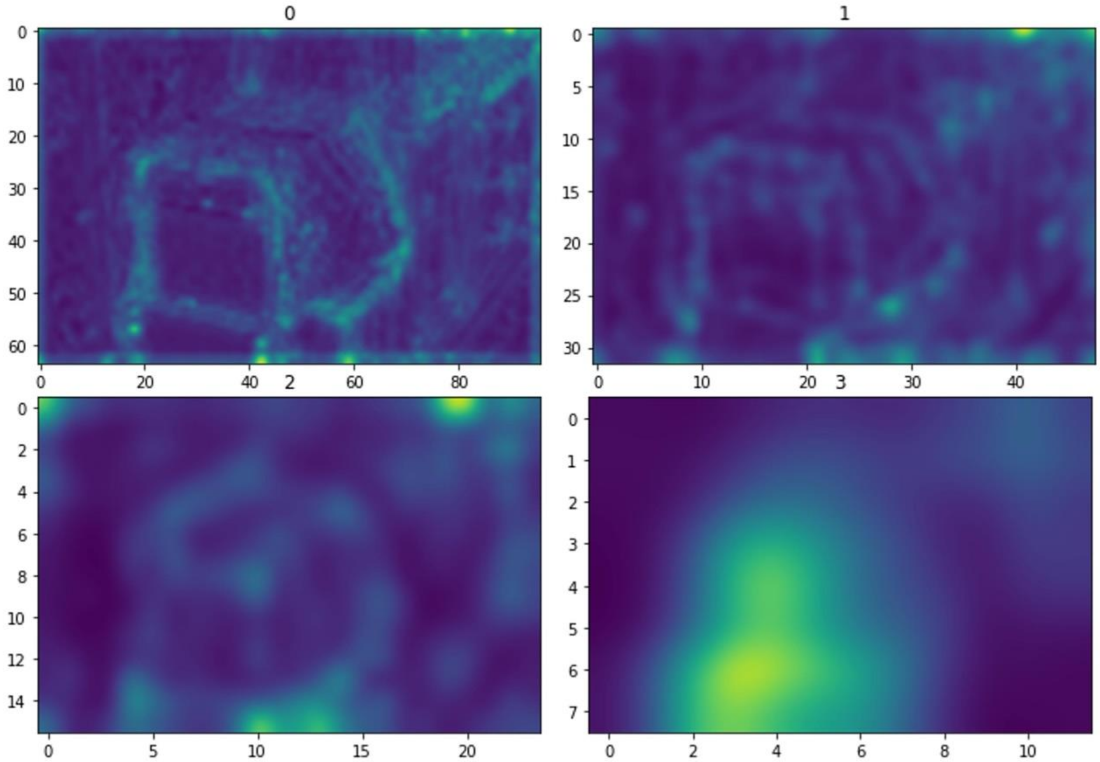

Figure 3: The attention map outputs the results among the layers of the residual block.

2.1.2. Overall structure of the network

The overall structure of the low illumination image enhancement network model based on knowledge distillation is shown in Figure 4. The network model is divided into two parts: Net-T and Net-S (image enhancement network). Net-T adopts the automatic encoding and decoding structure, and Net-S adopts the network structure similar to Net-T. Net-S coding part adopts partial hybrid dilated convolution.

ISSN 2616-5767 Vol.4, Issue 2: 52-56, DOI: 10.25236/AJETS.2021.040208

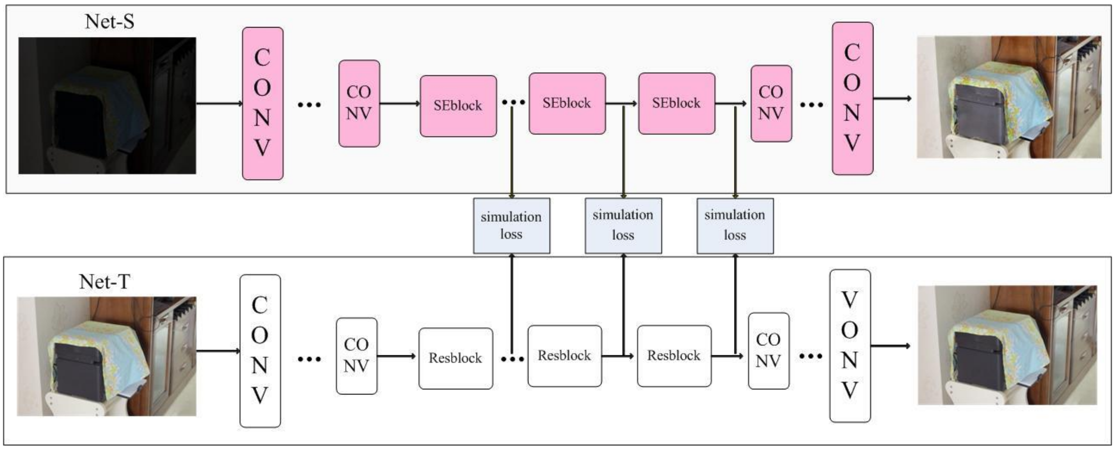

Figure 4: Overall structure of mine low illumination image enhancement network model based on knowledge distillation.

2.2. Experimental verification

After the experimental test, the real scene images of the mobile robot moving in the low illumination area of the surveillance video are randomly selected. The specific process is to extract the low illumination images in the video frame to form a test set, and randomly select the sample images for enhancement test. The resolution and size of image input are the same, without any advanced image processing. In the real scene, there is no normal illumination contrast image, and the enhancement effect of a single non contrast image is evaluated by using the NIQE index. The image of low illumination environment in coal mine is enhanced. The experimental analysis shows that the proposed algorithm has strong robustness, and the comprehensive performance index is improved by 42.21%, 7.47% and 30.18% respectively compared with CLAHE and LIME algorithm. Experimental results show that this method can significantly improve the image contrast and overall brightness, while reducing noise and more in line with human visual perception, which can meet the needs of intelligent mine video monitoring system construction.

2.2.1. Experiment introduction

In order to verify the effectiveness and superiority of this image enhancement method in low illumination image enhancement of mine video monitoring, the low illumination image of real mine monitoring video is selected for experiment, and the test image of data set is analyzed.

The ablation experiment is used to verify the network structure and loss function, and the image enhancement effect under different network structure and loss function is compared. Subjective visual perception and objective indicators are used to evaluate the contrast enhancement effect and objective performance of the proposed method and its three low illumination image enhancement methods. Among them, the three contrast methods are divided into two categories. One is the enhancement method without deep learning, including CLAHE and live. The other is based on a neural network. The real scene test image of the experiment comes from the monitoring video of the underground mobile robot on the roadway condition, and the image of the low light area moving in the monitoring video is randomly selected.

2.2.2. Experimental result

The real scene images of low illumination area in the mobile surveillance video of mobile robot are randomly selected for testing. In the contrast experiment, the preset parameters are used. Under the premise of the same output image size and resolution, the random sample images are respectively input for testing. The enhancement effect is as shown in the figure. The analysis of the processing results shows that the image quality can be improved by using CLAHE and LIME methods However, there are still some problems in the enhancement results. In Experiment 3, random images in surveillance video were enhanced respectively. From the perspective of enhancement effect, CLAHE and LIME enhanced images contain a lot of noise, the overall texture is fuzzy, and the details are not real. Some areas with high brightness are over enhanced in the live method, which affects the normal visual perception, enhances the brightness as a whole, and contains a lot of noise. The overall improvement of brightness by CLAHE method is not obvious. To a certain extent, this method overcomes some shortcomings of the above

ISSN 2616-5767 Vol.4, Issue 2: 52-56, DOI: 10.25236/AJETS.2021.040208

methods, improves the contrast and brightness of the enhanced image obviously, and contains less noise. While ensuring the image details and overall brightness, it is more in line with the human visual perception.

3. Conclusions

The network proposed in this paper is an end-to-end framework for low illumination image enhancement in coal mine. The experimental analysis shows that the idea of knowledge distillation method is applied to the low illumination image enhancement processing under the mine, and the hidden feature information in the net-t training process is extracted to guide the training of the enhancement network, which effectively increases the network operation efficiency and improves the effect of image enhancement. The ablation experiments verify that the pre trained net-t is more effective than the unused and randomly initialized Net-T in guiding the enhancement network, and verify the effectiveness of the influence of the components of the multi-scale loss function on effect of the image enhancement. The image of low illumination environment in coal mine is enhanced. The experimental analysis shows that the proposed algorithm has strong robustness, and the comprehensive performance index is improved by 42.21%, 7.47% and 30.18% respectively compared with CLAHE and LIME algorithms. Experimental results show that the proposed method can significantly improve the image contrast and overall brightness, while reducing noise and more in line with human visual perception, to a certain extent to meet the needs of intelligent mine video monitoring system construction.

References

[1] YAO H, WANG H, LI Y, et al. Three-dimensional spatial and temporal distributions of dust in roadway tunneling [J]. International Journal of Coal Science & Technology, 2020, 7(1): 88-96.

[2] PIZER S M, AMBURN E P, AUSTIN J D, et al. Adaptive histogram equalization and its variations [J]. Computer Vision, Graphics, and Image Processing, 1987, 39(3): 355-68.

[3] Suo J, Zhu S C, Shan S, et al. A compositional and dynamic model for face aging[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 32(3): 385-401.

[4] Zeng J, Ma X, Zhou K. CAAE++: Improved CAAE for age progression/regression[J]. IEEE Access, 2018, 6: 66715-66722.

[5] Lample G, Zeghidour N, Usunier N, et al. Fader networks: Manipulating images by sliding attributes[A]. Advances in Neural Information Processing Systems[C]. 2017: 5967-5976.

[6] XUAN D, GUAN W, YI P, et al. Fast efficient algorithm for enhancement of low lighting video; proceedings of the 2011 IEEE International Conference on Multimedia and Expo, F 11-15 July 2011, 2011 [C].

[7] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. The journal of machine learning research, 2014, 15(1): 1929-1958.

[8] BREIMAN LEO. Two-Eyed Algorithms and Problems [M]. 2003: 9.

[9] ISOLA PHILLIP, ZHU JUN-YAN, ZHOU TINGHUI, et al. Image-to-Image Translation with Conditional Adversarial Networks [M]. 2017: 5967-76.