Improved AODNet for Fast Image Dehazing

Shiyu Chen[1](✉), Shumin Liu[1], Xingfeng Chen[1,2], Jiannan Dan[1], and Bingbing Wu[1]

[1] School of Software Engineering, Jiangxi University of Science and Technology, Nanchang 330013, China
6720210718@mail.jxust.edu.cn

[2] Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China

Abstract. Application scenarios such as unmanned driving and UAV reconnaissance have the requirements of high performance, low delay and small space occupation. Images taken in foggy days are easy to be affected by fog or haze, thus losing some important information. The purpose of image dehazing is to remove the influence of fog on image quality, which is of great significance to assist in solving high-level vision tasks. Aiming at the shortcomings of the current defogging method, such as slow defogging speed and poor defogging effect, this paper introduces the idea of FPCNet and the attention mechanism module, and proposes an improved AODNet fast defogging algorithm to ensure the defogging speed and defogging performance. The public dataset RESIDE was used for training and testing. Experimental results show that in terms of dehazing performance, the proposed algorithm achieves 25.78 and 0.992 in PSNR and SSIM respectively. In terms of dehazing speed, the proposed method is close to AODNet, with only 5 times more parameters than AODNet, but more than 100 times smaller than other methods.

Keywords: Fast image dehazing · AODNet · Attention mechanism · FPCNet

1 Introduction

In atmospheric fog or haze environment, large particulate matter absorbs object light and scatters atmospheric light, resulting in visual blur and low image quality, which reduces the processing performance of advanced computer vision tasks (image recognition, target detection, image segmentation, etc.). Therefore, it is of great significance to solve the visual blur in fog in the field of image processing. At present, according to the image dehazing technology, it can be divided into two categories: traditional dehazing methods and deep learning-based dehazing methods [1].

Traditional defogging methods include image enhancement based defogging methods and atmospheric model based image defogging [2]. Image dehazing methods based on image enhancement achieve dehazing through image denoising, improving brightness, contrast and other image indicators, mainly including histogram equalization [3],

C. Wu et al. (Eds.): MONAMI 2023, LNICST 559, pp. 154–165, 2024.
https://doi.org/10.1007/978-3-031-55471-1_12

Retinex algorithm [4], wavelet transform [5], homomorphic filtering [6] and other methods. The main principle of image dehazing methods based on atmospheric model is that the fog map model is constructed based on the atmospheric scattering mechanism, and the haze-free image is solved by statistically estimating the relevant parameter values. The main methods include dark channel dehazing algorithm, dark channel dehazing algorithm based on guided filtering, Bayesian dehazing algorithm, etc. Fattal et al. [7] achieved image dehazing by estimating atmospheric transmittance. He et al. [8] found that in the dark channel map data of more than 5000 haze-free images, about 75% of the pixels had a value of 0, and 90% of the pixels had a very low value, which was concentrated in, and thus proposed the dark channel prior theory. Meng et al. [9] regarded the dehazing problem as an optimization problem based on boundary contrast and regularization, and realized image dehazing based on this theory. The color of a haze-free image is approximately divided into hundreds of different colors, which can form a close cluster in RGB three-dimensional space, and the pixels contained in the cluster are distributed on the entire image plane. In the case of fog, different camera distances can be converted into different transmission coefficients. Therefore, Berman et al. [10] proposed a dehazing method. Zhong et al. [11] used support vector regression to learn a regression model so that it could accurately estimate the transmission map of hazy images. To solve the delay problem, Tarel et al. [12] proposed a dehazing method that makes its complexity linearly related to the number of image pixels. In addition, Zhang Di et al. [13] used the improved dark channel method to realize atmospheric light estimation, but the dehazing performance could not be guaranteed.

In recent years, deep learning technology has been applied to the field of image dehazing. The dehazing methods based on deep learning can be divided into indirect deep learning dehazing methods and direct deep learning dehazing methods. Among them, the indirect deep learning dehazing method constructs an estimation model based on the deep network, so that the network model learns various atmospheric parameters in the training samples, and inverts the haze-free image based on the atmospheric parameters. The basic principle of direct deep learning dehazing method is that the model is separated from the atmospheric scattering model through large sample learning. The input value is the fog map, and the output value is the dehazing map.

Cai et al. [14] constructed a DehazeNet network through convolutional neural network to learn the foggy images and predict the foggy media transmission map. Li et al. [15] proposed an end-to-end dehazing method to avoid large image dehazing errors caused by intermediate parameter estimation. On this basis, to realize image dehazing, Ren et al. [16] proposed a multi-scale convolutional neural network. Ju Qingqing et al. [17] constructed three convolution kernels of different scales to convolve the hazy image and obtain a rough transmittance map through feature learning of different scales, but the dehazing delay of this method is large. Mei et al. [18] proposed an end-to-end deep learning dehazing method based on the U-net model, which was composed of an encoding block, a feature extraction block and a decoding block. Qin et al. [19] proposed an end-to-end feature fusion attention network, which is composed of a basic block and a feature attention module, and the basic block is composed of a multi-level residual model and an attention module. In order to achieve effective feature extraction, the direct dehazing method based on deep learning needs to build a deeper network to ensure the image dehazing performance, but it will bring problems such as complex model, huge parameters and high calculation delay.

In summary, the current main dehazing methods only focus on the dehazing performance, and ignore the model footprint, dehazing speed and other indicators. Especially in the field of unmanned driving, UAV reconnaissance and other fields that rely on computer vision, the landing of such technology not only requires the algorithm to overcome all kinds of complex and adverse weather, but also needs to meet the advantages of low delay and small space occupation. Therefore, it is of great significance to propose an image dehazing algorithm with good dehazing performance, low dehazing delay and small space occupation.

In order to better balance the two aspects of dehazing speed and dehazing performance, this paper introduces the idea of FPCNet [20] and the attention mechanism module, and proposes an improved AODNet fast dehazing model, which improves the expression ability of the network and reduces the complexity of the network, and enhances the network's learning of the depth of field fog characteristics to a certain extent. Under the condition of ensuring the dehazing speed of AODNet, the performance of dehazing is improved.

2 AODNet Image Dehazing

2.1 Atmospheric Scattering Model

In foggy scenes, due to atmospheric occlusion, interference from large particulate matter and other reasons, the light of interfering objects enters the camera system through reflection and refraction, resulting in blurred vision and incomplete image information. The formation process of fog map is shown in Fig. 1.

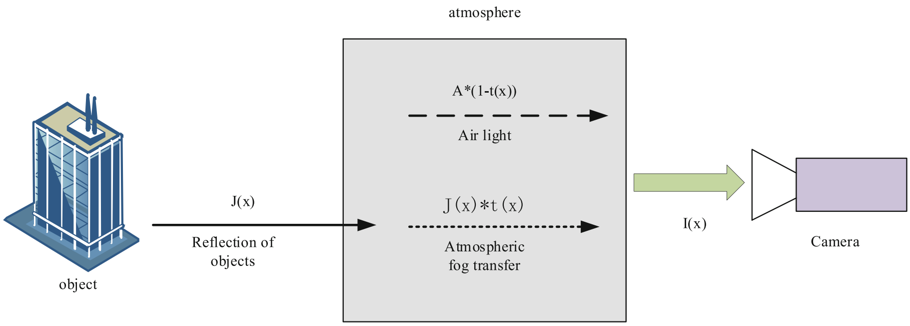

Fig. 1. Schematic diagram of the atmospheric scattering model

In Fig. 1, $I(x)$ is the hazy image taken by the camera, and $J(x)$ is the haze-free image of the corresponding scene. Based on the formation mechanism of atmospheric fog, the construction model of hazy image can be expressed as follows.

$$I(x) = J(x)t(x) + A(1 - t(x)) \tag{1}$$

Where, A is the atmospheric light value, $t(x)$ is the atmospheric transmittance, and $t(x)$ is defined as, $t(x) = e^{-\beta d(x)}$, β is the scattering coefficient, and $d(x)$ represents the depth information of the pixel.

According to the atmospheric light value A and transmittance $t(x)$, the haze-free image $I(x)$ can be inverted from the hazy image $J(x)$, as shown in Eq. 2:

$$J(x) = \frac{I(x) - A}{t(x)} + A \tag{2}$$

Equation 2 shows that the atmospheric light value and transmittance can be obtained by solving the haze-free image requirements. However, the obtained actual fog image cannot accurately estimate the value to obtain the optimal defogging effect, and the corresponding value can only be estimated by the atmospheric model and inversion to obtain the defogging image.

2.2 AODNet Dehazing

The dehazing algorithm based on convolutional neural network constructed by AODNet can realize end-to-end image dehazing. By estimating the atmospheric light value and transmittance at one time, that is, replacing A and t(x) with variable K(x) as a whole:

$$K(x) = \frac{(I(x) - A)/t + (A - 1)}{I(x) - 1} \tag{3}$$

Thus, the fog image model shown in Eq. (3) can be re-expressed as follows:

$$J(x) = K(x)I(x) - K(x) + 1 \tag{4}$$

The integrated variables $K(x)$ replace the transmittance and atmospheric light values after making the dependence on the fog map I(x) only. Based on the above single functional relationship, it can be obtained $K(x)$ by optimizing each weight parameter in the network through continuous iteration of deep learning, avoiding the empirical evaluation of transmittance and atmospheric light value. Compared with the traditional defogging methods, on the one hand, by constructing a convolutional neural network, the haze-free image is used as a label for training and learning, and then the value $K(x)$ is predicted and tends to the optimal value, so as to realize the image defogging function.On the other hand, simplifying the estimation of two different values into one value estimation can reduce the error accumulation, simplify the overall process, greatly reduce the processing delay compared with the traditional method, and improve the dehazing speed to a certain extent (Fig. 2).

Compared with the existing deep learning defogging algorithms, AODNet's defogging speed is a highlight. The design idea of network model and the core idea of defogging based on atmospheric model are worthy of reference. However, AODNet has the disadvantage of poor dehazing effect in removing fog images such as dark fog and thick fog [21]. Figure 3 shows the effect of AODNet to remove dark fog and thick fog, respectively.

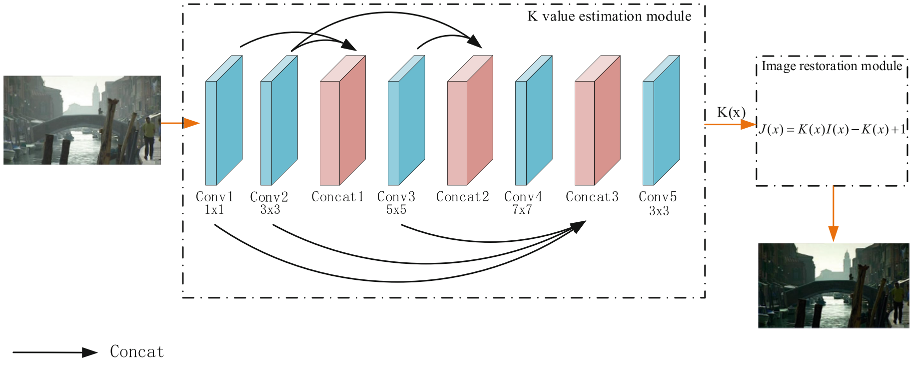

Fig. 2. Schematic diagram of AODNet network model

Fig. 3. Comparison of AODNet dehazing effect

3 Improved AODNet for Dehazing

Aiming at the problems of incomplete dehazing and loss of detailed information in AODNet dehazing algorithm, the idea and attention mechanism module of FPCNet are introduced into the dehazing algorithm proposed in this paper. Figure 4 is the overall network architecture of this paper. The network consists of two parts, k value estimation module and image restoration module. The part to the left of AB in the k value estimation module is called the feature extraction module. The feature extraction module is improved on the basis of the structure of the original AOD, and consists of four 1x1 convolution operations, three pooling operations of different sizes, and three feature concatenation operations. The AB and the right part of the k value estimation module are called feature enhancement module, which is composed of an attention mechanism and two layers of convolution. The attention mechanism can make the model realize self-attention of image features when dealing with local image information such as defogging brightness and dark, so as to solve problems such as dark local defogging. Finally, through the image restoration module, the value of k estimated by the model was substituted into the atmospheric scattering model to obtain the clear image J(x).

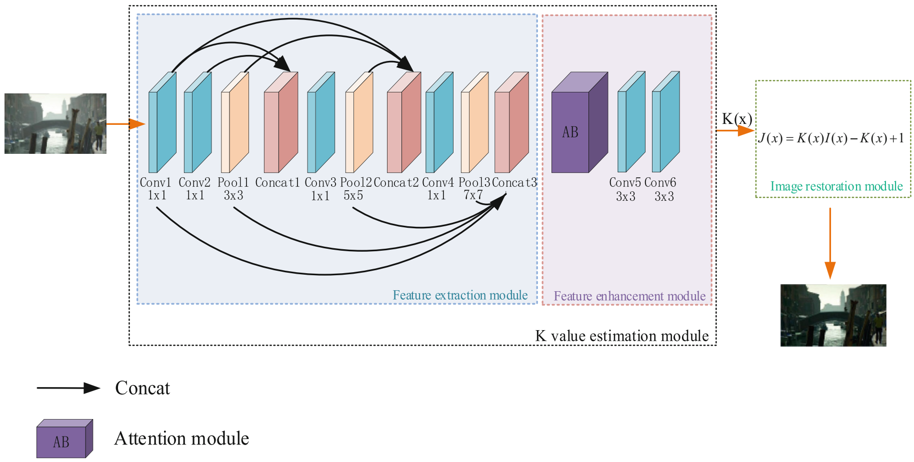

Fig. 4. Schematic of the overall network model in this paper

3.1 Feature Extraction Module

In order to improve the feature representation ability of AOD-Net, inspired by the full pointwise convolution unit of FPC-Net, this paper uses pointwise convolution and different sizes of pooling to replace the large-scale convolution in AODNet. This improved method can not only effectively improve the expression ability of the network for features, make the network more compact, but also reduce the overfitting phenomenon during training to a certain extent [22]. In addition, this paper believes that it is necessary to fuse the features of Conv1 into the Concat2 fusion layer. It is found through experiments that such optimization can improve the network's ability to process the depth of field fog to a certain extent.

The network structure of the feature extraction module consists of four 1x1 convolution operations, three pooling operations of different sizes, and three feature concatenation operations. Among them, the size of the convolution kernel of each pointwise convolution is 1×1, the 1×1 pointwise convolution layer of the input layer is followed by no pooling layer and batch normalization layer, and the remaining pointwise convolution will connect the batch normalization layer, ReLU activation function and pooling layer. The first pooling layer has a kernel size of 3×3 and a padding value of 1, the second pooling layer has a kernel size of 5×5 and a padding value of 2, and the third pooling layer has a kernel size of 7×7 and a padding value of 3 with a step size of 1 for each convolution kernel and pooling layer. Concat1 is concatenated from conv1, conv2, concat2 is concatenated from conv1, pool1, pool2, and concat3 is concatenated from conv1, pool1, pool2, and pool3. The details of the network parameters of the feature extraction module are shown in Table 1.

Table 1. Network parameters of feature extraction module

	Input Size	Filter	Filter Number	Pad	Stride
Conv1	480 × 640 × 3	1 × 1	32	0	1
Conv2	480 × 640 × 32	1 × 1	32	0	1
Pool1	480 × 640 × 32	3 × 3	—	1	1
Concat1	480 × 640 × 64	—	—	—	—
Conv3	480 × 640 × 32	1 × 1	32	0	1
Pool2	480 × 640 × 32	5 × 5	—	2	1
Concat2	480 × 640 × 96	—	—	—	—
Conv4	480 × 640 × 32	1 × 1	32	0	1
Pool3	480 × 640 × 32	7 × 7	—	3	1
Concat3	480 × 640 × 128	—	—	—	—

3.2 Attention Mechanism Module

Image quality and details may be affected to different degrees when performing operations such as object classification, compression, and restoration on images, which may affect subsequent high-level vision tasks. In order to overcome the influence of image quality differences on the dehazing effect, this paper adds an attention mechanism module based on the improved AODNet model, so as to solve the shortcomings of dark image dehazing and low generalization of the model limited by samples. The fusion of channel attention mechanism and spatial attention mechanism enables the model to realize self-attention of image features when dealing with local image information such as defogging bright and dark, and ensures that the model solves problems such as local dark defogging (Fig. 5).

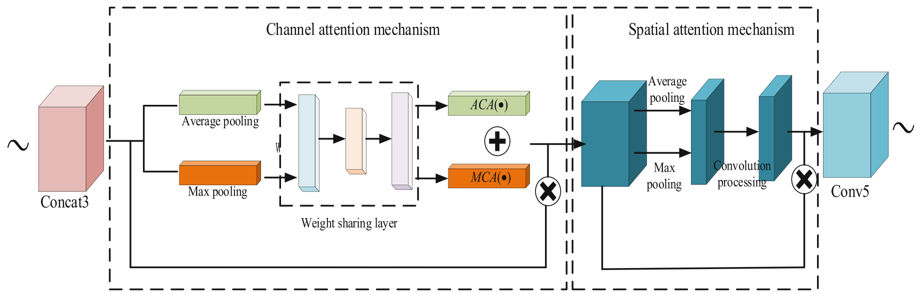

Fig. 5. Schematic of the attention mechanism module

The attention module adopted in this paper is jointly composed of a channel attention module and a spatial attention module. Firstly, the feature maps obtained by concat3 were processed by average pooling and maximum pooling respectively to realize the channel feature map processing from 3D feature map to 2D. Secondly, the channel feature maps after two processing are processed by a 1* 1 convolution kernel to realize adaptive channel weight calculation. The corresponding positions of the channel weights obtained by the two are added together, which is expressed as:

$$AMCA(K) = ACA(K) + MCA(K) \tag{5}$$

Thirdly, the added weights were normalized by the sigmoid function. Finally, the spatial attention weights are multiplied with the input feature map K to achieve the same dimension, and the channel self-attention is expressed as follows:

$$CA(K) = K * sig\text{mod}(AMCA(K)) \tag{6}$$

The feature maps $CA(K)$ obtained after processing by the channel attention mechanism are firstly used as the input values of the spatial attention module to undergo average pooling and maximum pooling operations respectively. Secondly, the results of the two pooling were concatenated together. Thirdly, the sigmoid function was used to normalize the weights to ensure that the weights were within a reasonable range. Finally, the weights are multiplied with the input values $CA(K)$ to achieve the corresponding position, and the feature map is output to conv5.

4 Experiment and Analysis

4.1 Experimental Scheme

Dataset. In this paper, the public Reside- β dataset is used as the data source. The training set is composed of 2061 clear outdoor images each generating 35 types of synthetic fog images, and the synthetic fog atmospheric light values A = [0.8,0.85,0.9,0.95, 1], β = [0.04,0.06,0.08,0. 1,0. 12,0.16,0.2] are obtained. A total of 72135 synthetic fog images are obtained. The test Set is selected from the outdoor data set of OTS (Objective Testing Set), which consists of 500 images.

Experimental Environment. The experimental environment of this paper is as follows: Win10 operating system, processor InterlCore i5-12400F CPU, GTX 3060TI GPU accelerated computing, python3.8 programming language, and Pytorch deep learning framework.

Parameter Setting. The clear image is taken as the true label, the MSE defines the model loss function, the Gaussian random variable is used to initialize the weight value of the network model, the initial learning rate is set to 0.0001, and the momentum and decay are set to. The Adam optimizer was used to realize the loss convergence to minimize, and the weight update optimization model of the network was trained, with a total of 50 iterations.

4.2 Comparative Analysis of Dehazing Performance

Because the synthetic fog image is synthesized from the clear image through the atmospheric scattering model, the corresponding clear image can be found when evaluating the defogging performance of the synthetic fog image. PSNR and SSIM are used as the classical image quality indicators to quantify the defogging performance of various algorithms, and the higher the value is, the better the image quality is. PSNR indicates how close the dehazed image is to the original image, and SSIM describes the structural similarity between the dehazed image and the haze-free image.

Table 2 lists the index results of the synthetic fog data set OTS (Outdoor) tested by each method. The data PSNR and SSIM in the table are obtained by taking the average value last. The algorithm in this paper performs better than other algorithms, reaching 25.78 and 0.992. The proposed algorithm improves the feature extraction ability of the original AODNet, so that the image restoration effect is more excellent.

As shown in Fig. 6, some dehazing maps are selected from OTS (Outdoor) test results for subjective evaluation. DCP dehazing is prone to color distortion and halo phenomenon, and the overall brightness is dark. Although the overall dehazing effect of FFA is good, the effect is often not very impressive in the case of low image contrast. Some of the results of PFF dehazing appear dark shadows, and there are a lot of fog residue. GCA needs to be improved in removing dark fog, and the color reproduction is not high. After AOD dehazing, there are some fog residues in the image, and the image details cannot be restored well. The defogging images obtained by the proposed algorithm have high color fidelity, obvious defogging effect, and satisfactory defogging quality.

Table 2. Comparison of synthetic fog dehazing performance of different algorithms

	PSNR	SSIM
DCP	16.57	0.941
FFANet	23.34	0.989
GCANet	20.94	0.960
PFFNet	19.96	0.982
AODNet	22.97	0.976
Our model	25.78	0.992

Fig. 6. Comparison of defogging effects of different algorithms

4.3 Comparative Analysis of Dehazing Speed

In order to compare the dehazing speed of various dehazing algorithms, the dehazing speed of each algorithm is compared in the same experimental environment, and the time consumed by various dehazing algorithms to process 500 fog images of OTS dataset is counted. Table 3 lists the average dehazing time of a single image. Since the DCP dehazing method is the traditional dehazing method, there is no model parameter value, and the deep learning models are all calculated by GPU acceleration under the same conditions, as shown in the table.

Table 3. Dehazing time for different dehazing algorithms

	time (s)	Programming languages	Number of parameters
DCP	0.0754	Python	X
FFANet	0.4892	Python	16.998 MB
GCANet	0.0573	Python	2.681 MB
PFFNet	0.0223	Python	57.038 MB
AODNet	0.0258	Python	0.0067 MB
Our model	0.0392	Python	0.0269 MB

According to the table, in terms of dehazing speed, PFFNet has the fastest dehazing speed, followed by AODNet and FFANet has the slowest dehazing speed. Because FFANet and GCANet both construct a deeper dehazing model, the dehazing speed is slow. The proposed model is only slightly lower than AODNet, and the average dehazing time of a single image is 0.0392 s. Although the number of parameters can not reflect the operation time of the model, it can reflect whether the model is more complex. In terms of the number of parameters, PFFNet, GCANet and FFANet construct more complex dehazing models, while AODNet constructs a lightweight dehazing model. For more complex dehazing models, the proposed model reduces the number of parameters by more than 100 times, and only increases the number of parameters by 4 times than the lightweight AODNet.

In general, the improved dehazing method in this paper achieves a good balance between time and performance, is close to AODNet in dehazing speed, and is superior to other dehazing methods in dehazing performance.

5 Summary

In order to solve the problems of large space consumption and slow defogging speed of the current defogging algorithm, this paper proposes an improved AODNet fast defogging network model based on the fast defogging idea of AODNet. The FPCNet idea is introduced to replace the convolution of AODNet with point-wise convolution and pooling operation. The features of the second feature concatenation layer and the first convolutional layer are fused to form a dense connection structure, and the attention mechanism is introduced to make up for the shortcomings of AODNet in dehazing details. The experimental results on OTS synthetic fog data set show that the proposed dehazing algorithm is superior to the traditional algorithm and deep learning algorithm in the classic image quality indicators PSNR and SSIM. The parameter amount is reduced by more than 100 times, and the dehazing speed can reach an average of 0.0392 s per image. A good balance among performance, space and time is achieved.

References

1. Xie, Y., Jia, H., Wang, T., Lei, C., Xu, K., Chen, Q.: Review of image dehazing algorithms∗. Comput. Dig. Eng. **50**(12), 2765–2774 (2022)
2. Wu, D., Zhu, Q.: Research progress of image dehazing. Acta Automatica Sinica **41**(2), 221–239 (2015)
3. Jiang, H., Yang, Z., Zhang, X.: Research progress of image dehazing algorithm. J. Jilin Univ. Eng. Technol. Ed. **51**(4), 1169–1181 (2021)
4. Li, X.: Image enhancement algorithm Based on Retinex theory. Comput. Appl. Res. **22**(2), 235–237 (2005)
5. Chen, S., Cao, S., Cui, M.Y.: Blind image deblurring algorithm based on deep multilevel wavelet transform. J. Electron. Inf. Technol. **43**(1), 1–8 (2021)
6. Ma, X.: Ship image enhancement method based on homomorphic filtering. Ship Sci. Technol. **42**(6), 89–91 (2020)
7. Fattel, R.: Single image dehazing. ACM Trans. Graph. **27**(3), 721–728 (2008)
8. He, K., Sun, J., Tang, X.O.: Single image haze removal using dark channel prior. IEEE Trans. Pattern Anal. Mach. Intell. **33**(12), 2341–2353 (2011)
9. Meng, G.F., Wang, Y., Duan, J.H., et al.: Efficient image dehazing with boundary constraint and contextual regularization. In: IEEE International Conference on Computer Vision, pp. 617–624 (2013)
10. Berman, D., Treibitz, T., Avidan, S.: Non-local image dehazing. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016)

11. Zhong, L., Shang, Y., Zhou, X., et al.: Fast single image dehazing based on a regression model. Neurocomputing **245**, 10–22 (2017)
12. Tarel, J.-P., Hautiere, N.: Fast visibility restoration from a single color or gray level image. In: IEEE 12th International Conference on Computer Vision, pp. 2201–2208 (2009)
13. Zhang, D., Wu, P.: Fast dehazing algorithm for single image. Comput. Eng. Appl. **55**(10), 213–217 (2019)
14. Cai, B., Xu, X., Jia, K., et al.: DehazeNet: an end-to-end system for single image haze removal. IEEE Trans. Image Process. **25**(11), 5187–5198 (2016)
15. Li, B., Peng, X., Wang, Z., Xu, J., Feng, D.: Aod-net: all-in-one dehazing network. In: IEEE International Conference on Computer Vision, pp. 4770–4778 (2017)
16. Ren, W., Liu, S., Zhang, H., et al.: Single image dehazing via multi-scale convolutional neural networks. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) Computer Vision-ECCV, pp. 154–169. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-319-46475-6_10
17. Ju, Q., Li, C., Sang, Q.: Single image dehazing method based on improved multi-scale convolutional Neural Network. Comput. Eng. Appl. **55**(10), 179–185 (2019)
18. Mei, K., Jiang, A., Li, J., et al.: Progressive feature fusion network for realistic image dehazing (2018)
19. Qin, X., Wang, Z., Bai, Y., et al.: FFA-net: feature fusion attention network for single image dehazing. In: Proceedings of the AAAI Conference on Artificial Intelligence (2020)
20. Zhang, J., Tao, D.: FAMED-net: a fast and accurate multi-scale end-to-end dehazing network. IEEE Trans. Image Process. **29**, 72–84 (2019)
21. Tran, L.-A., Moon, S., Park, D.-C.: A novel encoder-decoder network with guided transmission map for single image dehazing. Procedia Comput. Sci. **204**, 682–689 (2022)
22. Zhang, J., Cao, Y., Wang, Y., Zha, Z.J., Wen, C., Chen, C.W.: Fully point-wise convolutional neural network for modeling statistical regularities in natural images. ArXiv (2018)