A high-resolution image dehazing GAN model in icing meteorological environment

Xinling Yang[a], Wenjun Zhou*[a], Chenglin Zuo[b], Yifan Wang[a], Bo Peng[a]

[a]School of Computer Science, Southwest Petroleum University, Chengdu Sichuan, China;
[b]Key Laboratory of Icing and Anti/De-Icing, China Aerodynamics Research and Development Center, Mianyang Sichuan, China

ABSTRACT

In this paper, we propose a high-resolution GAN model for image dehazing in icing meteorological environment, which strictly follows a physics-driven scattering strategy. First of all, the utilization of sub-pixel convolution realizes the model to remove image artifacts and generate high-resolution images. Secondly, we use Patch-GAN in the discriminator to drive the generator to generate a haze-free image by capturing the details and local information of the image. Furthermore, to restore the texture information of the hazy image and reduce color distortion, the model is jointly trained by multiple loss functions. Experiments show the proposed method achieves advanced performance for image dehazing in icing weather environment.

Keywords: Generative adversarial networks, Image dehazing, Sub-pixel convolution, High resolution

1. INTRODUCTION

When an aircraft is flying at high altitude, contacting with water vapor in the air will cause icing on the surface of the aircraft, and the icing of the aircraft will greatly affect flight safety. During the test of this scene in the icing wind tunnel, there will be suspended water droplets with a certain liquid water content in the test section. When the light passes through the water droplets in the test section, the generated image will be blurred due to the attenuation of light scattering and absorption, resulting in a loss of image quality. These images usually have distorted colors, decreased contrast, lost edge and texture information, and cannot accurately observe the conditions of icy areas. Therefore, it is of practical significance to apply image dehazing to improve the quality of monitoring images in icing wind tunnel test.

In common dehazing algorithm, an atmospheric scattering model[1] is usually used to describe the relationship between the hazy image and the haze-free image:

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

where, $J(x)$ represents the haze-free image, $I(x)$ represents the collected hazy image, $t(x)$ represents the transmittance map, and A represents the atmospheric illumination value. The essence of image dehazing is the process of restoring the hazy image to the fogless image infinitely. After $I(x)$ is given, in order to find $J(x)$ for the reverse solution, we usually focus on finding the value of both $t(x)$ and A. Image dehazing is a highly ill-posed problem.

Many traditional algorithms for image dehazing focus on solving the transmittance and atmospheric illumination values in the atmospheric scattering model. Representative algorithms include dark channel prior dehazing algorithm DCP[2] and color attenuation prior dehazing algorithm CAP.[3] In recent years, with the rise of deep learning methods, algorithms for image dehazing have also achieved good results. Among them, there are also many dehaze maps that combine neural networks with prior knowledge, that is, estimate transmittance and atmospheric illumination values through network learning, and then obtain dehaze maps by inversely solving atmospheric scattering models, such as DehazeNet,[4] multi-scale depth dehaze network MSCNN,[5] dense pyramid dehazing algorithm DCPDN[6] and At-DH,[7] etc.

In icy weather conditions, the concentration of hazy is often relatively high. For this dense hazy situation, we learned the related method At-DH in the NTIRE19 dehazing challenge. The dehazing data set DenseHaze[8]

*Wenjun Zhou: E-mail: zhouwenjun@swpu.edu.cn

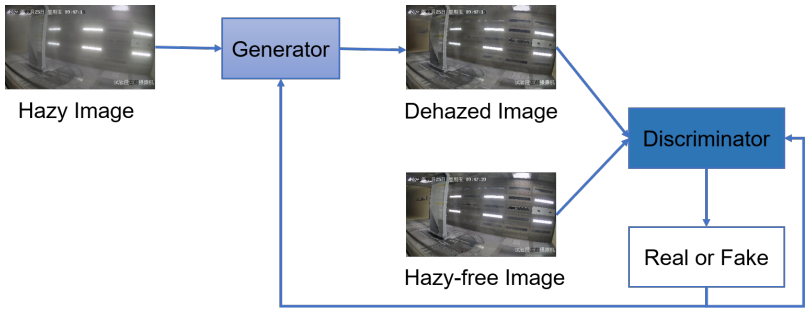

Figure 1. Network structure diagram.

in NTIRE19 is characterized by dense and uniform hazy, which are similar with the pictures obtained under the environment. The algorithm At-DH has achieved good results in this challenge, and the dehazing effect is obvious. Therefore, inspired by the algorithm At-DH, in this paper we propose a high-resolution GAN model with dense connection for image dehazing in icing meteorological environment. According to the inverse problem of image dehazing, we estimate the transmittance and atmospheric illumination value through network learning, reverse the atmospheric scattering model to perform the dehazing work, and strictly follow the physics-driven scattering model to achieve better dehazing effect. And the generation confrontation network GAN has great potential in the research direction of real image restoration. We combine this strategy with image dehazing to generate a haze-free image that is closer to the real one. The main contributions of this paper are as follows:

1. A high-resolution GAN model for image dehazing is proposed in this paper to obtain relatively dehazed images in icing meteorological environment.

2. The proposed model can eliminate the artifacts caused by traditional de-convolution to a certain extent and help to obtain high-resolution dehazed image.

2. RELATED WORK

Image dehazing algorithms are generally divided into prior knowledge-based and learning-based methods. The dehazing algorithm based on prior knowledge can be divided into physical model algorithm and non-physical model algorithm: the physical model dehazing algorithm is based on the atmospheric scattering model, and the related parameters $t(x)$ and A value are obtained through prior knowledge. The hazy-free image obtained by Equation (1), such as algorithms DCP, CAP, etc.; the non-physical model defogging algorithm uses image enhancement methods to dehaze, such as Retinex defogging algorithm,[9] histogram equalization,[10] wavelet and homomorphic filtering algorithms. At the same time, the learning-based dehazing algorithm can also be similarly divided into two parts, the estimated parameter method and the direct restoration method: the estimated parameter method is to estimate $t(x)$ and A through network learning to perform defogging, and the parameters estimated by using deep learning are generally more accurate than the traditional ones, such as MSCNN,[5] DehazeNet[4] and DCPDN;[6] the direct repair method is to directly learn and estimate the output dehaze image from the input fog image through the network, such as FD-GAN,[11] GridDehazeNet[12] and FFA-Net.[13]

3. METHODOLOGY

In this section, we will introduce our network structure in detail. The network is based on the generation confrontation network GAN, which consists of two parts: the generator and the discriminator. The atmospheric scattering model inversely solves the fog-free image, and the discriminator distinguishes the hazy-free image and its corresponding real hazy-free image. The whole framework is shown in Fig. 1.

3.1 Generator

The specific network structure diagram of the generator is shown in Fig. 2. The main function of the encoder on the left is to extract important features of the picture, while the decoder on the right estimates scene information

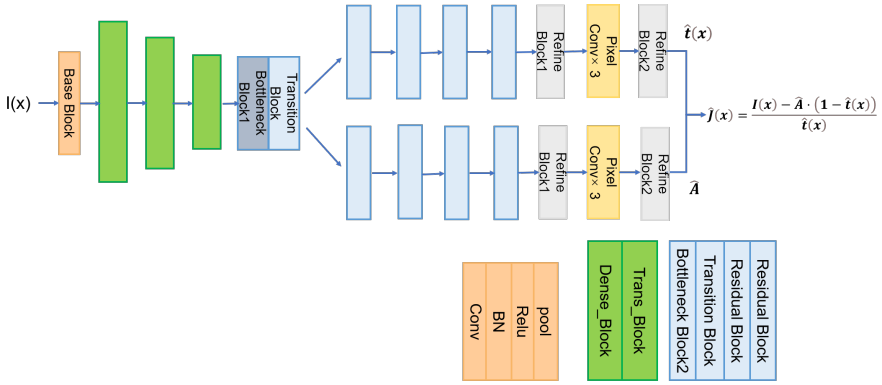

Figure 2. Generator. Learning to estimate parameters $t(x)$ and A through densely connected networks.

based on the features extracted from the encoder, and at the same time restores the original size of the picture. The encoder is constructed based on the densely connected network DCN,[14] mainly the densely connected module (Dense_Block) and the transmission module (Trans_Block). The structure of the decoder is similar to that of the encoder, including the bottleneck layer(Bottleneck Block),[7] the transmission layer (Transition Block), the residual layer (Residual Block) and the refinement layer (Refine Block).[6]

3.1.1 Encoder

In the network structure, the encoder's BaseBlock to the third Trans_Block use the pre-training parameters of the first half of the DCN, because the structure in the network realizes the splicing of features on the channel, thereby achieving the effect of feature reuse. Using pre-trained connected blocks in the dehazing work helps to obtain important features.

3.1.2 Decoder

The main function of Transition Block is to enlarge and change the refinement features. The channel change is completed by the 1×1 convolutional layer in the Tranition Block, and then the feature is enlarged through upsampling. Compared with Dens_Block in the encoder, Bottleneck Block adds batch normalization once to normalize the training data, so that the network has better training stability and avoids gradient explosion. Adding a residual layer between two consecutive dense blocks helps recover more details in the image by extracting more high-frequency information.

Refine Block merges and retouches image information at different scales. The addition of the residual network enables superimposing the identity mapping layer on the shallow network so that the network does not degrade as the depth increases.

3.1.3 Sub-pixel convolution

Sub-pixel convolution[15] is applied in the field of image super-resolution, which can super-resolve low resolution data to high resolution space, and is an upsampling method of pixel rearrangement. The sub-pixel convolution process is described as follows:

$$I^{HR} = f^{L}(I^{LR}) = PS(W_L \times f^{L-1}(I^{LR}) + b_L). \tag{2}$$

I^{HR} is a high-resolution image, I^{LR} is a low-resolution image, f is a convolution operation, W_L is the weight of the convolution kernel, b_L is the bias item, and PS is the pixel reorganization operation.

The pixel reorganization operation is to take an element from each channel of the multi-channel feature map and combine it into a square unit on the new feature map. The pixels on the original feature map are equivalent to the sub-pixels on the new feature map. In this paper, we replace the original up-sampling layer with the sub-pixel convolution layer, so that there is no need to add meaningless 0 elements during the up sampling process, and to a certain extent eliminate the artifacts caused by traditional inverse convolution, while helping to reduce computational complexity.

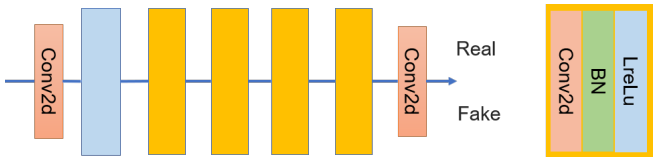

Figure 3. Discriminator. It is beneficial to drive the generator to capture the local information of the image and generate a relatively high-resolution haze-free image.

3.2 Discriminator

Here the discriminator introduces PatchGAN.[16] Compared with the usual discriminator, it discriminates the image by sub-area matrix, and finally takes the average result of all the matrices to output true and false. It can pay more attention to the image details when training the model, and obtains higher resolution pictures. Since Patch-GAN has fewer parameters and runs fast, it can be applied to pictures of any size. The discriminator consists of a series of convolutional layers, batch normalization layers, and activation layers, where as shown in Fig. 3.

3.3 Loss function

The loss function is used to standardize the learning direction of parameters in network training. In order to better train the network and generate images with good dehazing effect, we use three common losses: reconstruction loss, perception loss and confrontation loss.

3.3.1 Reconstruction loss

We use the reconstruction loss to compare the gap between the generated dehazed image and the real haze-free image in image pixel space, which can be expressed as:

$$L_{Res} = \frac{1}{N}\sum_{i=1}^{N} ||G(I_i) - J_i||_1, \tag{3}$$

where I_i represents the input foggy image, J_i represents the real fog-free image corresponding to the image, and $G(I_i)$ represents the dehazing generated by the generator picture.

3.3.2 Perceptual loss

We also use perceptual loss here to measure the perceptual similarity in the feature space of the dehazed image and the haze-free image. The specific implementation is to evaluate the parameters of the VGG16 pre-trained network model as follows:

$$L_p = \frac{1}{N}\sum_{i=1}^{N} ||\phi(G(I_i)) - \phi(J_i)||_2^2, \tag{4}$$

where $\phi(.)$ represents the feature map obtained from the VGG16 network layer.

3.3.3 Adversarial loss

In the generated confrontation network, in order to restore the authenticity of the image, the confrontation loss is the most commonly used type of loss as follows, and the binary cross-entropy function is used to calculate the loss value.

$$L_A = \frac{1}{N}\sum_{i=1}^{N} \log\left(1 - D\left(J_i, G\left(I_i\right)\right)\right), \tag{5}$$

3.3.4 Overall loss function

Finally, the overall loss function consists of the reconstruction loss function, perceptual loss function, and adversarial loss function. Its definition is as follows:

$$\mathrm{L}_{all} = L_{\mathrm{Res}} + \alpha_1 L_p + \alpha_2 L_A. \tag{6}$$

where α_1 and α_2 are the function weight.

4. EXPERIMENT

4.1 Settings

The hazy images involved in this paper are images taken from multiple angles in the icing wind tunnel experimental scene supported by the Key Laboratory of Icing and Anti/De-icing of CARDC. Some pictures are shown in Fig. 4. In order to simulate the state of the aircraft passing through the cloud layer containing super-cooled water droplets, cloud field with a certain water droplet diameter MVD (median volume diameter) and water content LWC (liquid water content) can be selected by adjusting the water pressure and air pressure of the nozzle during the icing wind tunnel test. Both MVD and LWC are important parameters for determining the haze of cloud field. The data sets we use here include foggy images with MVD of 25 μm and LWC of 1.31 g/m^3, MVD of 22 μm and LWC of 1.19 g/m^3, and MVD of 20 μm and LWC of 1.0 g/m^3 and 0.5 g/m^3 respectively.

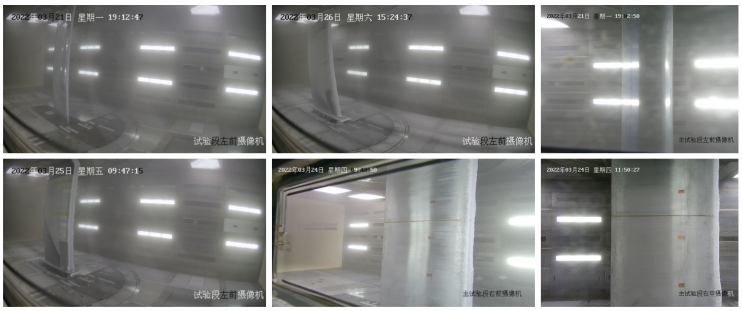

Figure 4. Some sample images from the icing wind tunnel.

In this paper, we selected 310 cropped foggy images in the icy wind tunnel as the training set and trained them on the GPU. During the training, the input image size will be resized to 1024×1024. The Adam optimizer is used for the generator and the discriminator, and the learning rate of both is set to 10^{-4}, and a total of 100 epochs are iterated. The experiment runs in the environment of P6000 GPU, 24GB memory to train the model.

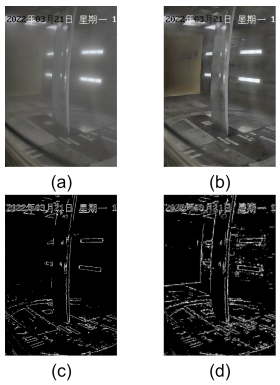

Figure 5. Comparison of visible edges before and after dehazing, e=0.88, $\overline{r}$=2.91 after dehazing. (a) Original haze image; (b) Dehazed image; (c) Visible edge of original hazy image (d) Visible edge of image after dehazing.

4.2 Evaluation metrics

Since there is no corresponding real fog-free image, in our experiment, the widely used visible edge gradient method[17] was used to evaluate the experimental results. There are two important indicators: the ratio of the number of visible edges e and the average value of the regularized visible edge gradient $\overline{r}$. As the example shown in Fig. 5, after defogging, the overall contrast of the image is enhanced, and the number of visible edges measured increases.

The relevant expressions are as follows

$$e = \frac{n_r - n_0}{n_0}, \tag{7}$$

n_0 and n_r represent the number of visible edges before and after image defogging, respectively, and e represents the ability of the algorithm to restore invisible edges in the image. The larger the value, the better the defogging effect.

$$\overline{r} = exp(\frac{1}{n_r} \sum_{P_i \in Y_r} \log r_i) \tag{8}$$

Y_r is the set of pixels on the visible edge of the image after dehazing, r_i is the ratio of the gradient of the dehazing image and the foggy image at pixel point P_i. The larger the value of $\overline{r}$, the higher the contrast of the image after defogging, and the better the effect.

4.3 Comparative experiments

To evaluate our proposed model more accurately, a comparison with state-of-the-art dehazing methods is performed on icing wind tunnel haze images. Advanced methods include traditional dehazing methods: DCP,[2] CAP,[3] AMEDF,[18] LBF;[19] deep learning methods: AOD-NET,[20] MSCNN,[5] FFA-Net,[13] D4.[21]

Table 1 and Fig. 6 respectively shows the comparison results and visual effects of our model and traditional dehazing methods in the wind tunnel dataset. From Fig. 6, we can intuitively see that DCP, CAP, and LBF have a certain defogging effect and increase the number of visible edges, but the results are distorted and the overall color is dark. AMEF has a better defogging effect, especially in the case of haze with an LWC of 0.5. Overall, compared with these traditional methods, our model dehazes more thoroughly and has better visual effects. From Table 1, in the case of dense fog with LWC of 1.31 and 1.0, $\overline{r}$ can achieve the best result.

Table 1. Comparison of the results of traditional methods on the test set.

Metric	Image	DCP[2]	CAP[3]	AMEDF[18]	LBF[19]	**Ours**
	Image 1	1.67	3.51	1.59	1.81	**0.88**
	Image 2	6.73	14.64	4.54	11.80	**1.45**
e	Image 3	1.33	3.36	1.41	2.52	**0.82**
	Image 4	1.21	2.35	1.18	1.80	**0.72**
	Image 5	2.60	3.12	2.41	4.41	**0.44**
	Image 6	3.74	4.98	2.48	4.41	**0.63**
Average		2.88	5.33	2.27	4.46	**0.82**
	Image 1	1.22	1.21	2.35	1.19	**2.91**
	Image 2	1.16	1.46	2.48	1.84	**3.38**
$\overline{r}$	Image 3	1.04	1.17	2.50	1.24	**2.56**
	Image 4	1.13	1.11	2.49	1.17	**2.60**
	Image 5	1.34	1.50	2.06	1.94	**2.02**
	Image 6	1.24	1.24	2.23	1.47	**1.83**
Average		1.19	1.28	2.35	1.48	**2.55**

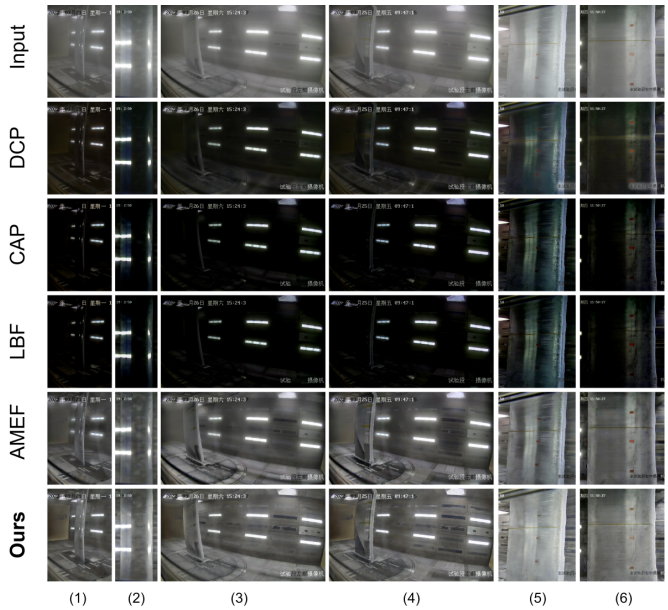

Figure 6. Comparison of six real hazy images in an icing wind tunnel with existing traditional methods. (1) and (2): LWC=1.31; (3) and (4): LWC=1.0; (5) and (6): LWC=0.5.

Table 2 shows the comparison results between our model and the deep learning defogging method in the wind tunnel dataset. For the deep learning method, we use the provided pre-trained model for the defogging test. The visual effect comparison after defogging is shown in Fig. 7. From the results, it can be seen that the color of AOD-NET is distorted, the effect of MSCNN on defogging is not obvious, the effect of D4 on the close-range of the picture is better, and FFA-Net causes picture distortion. However, our model also restores the foreground part of the picture to a certain extent. After dehazing, the picture is clearer and retains the original outline.

Table 2. Comparison of results of deep learning methods on the test set.

Metric	Image	AOD-NET[20]	MSCNN[5]	FFA-Net[13]	D4[21]	**Ours**
	Image 1	2.38	0.46	1.30	1.77	**0.88**
	Image 2	5.41	0.76	1.40	5.47	**1.45**
e	Image 3	1.75	0.30	0.86	1.56	**0.82**
	Image 4	1.39	0.26	0.63	1.23	**0.72**
	Image 5	2.00	1.16	1.40	1.09	**0.44**
	Image 6	2.94	0.68	0.72	1.73	**0.63**
Average		2.65	0.60	1.05	2.14	**0.82**
	Image 1	1.70	1.22	1.34	1.49	**2.91**
	Image 2	2.18	1.29	1.45	1.60	**3.38**
$\overline{r}$	Image 3	1.73	1.20	1.43	1.53	**2.56**
	Image 4	1.69	1.19	1.35	1.52	**2.60**
	Image 5	1.74	1.37	1.48	1.44	**2.02**
	Image 6	1.53	1.15	1.26	1.36	**1.83**
Average		1.76	1.24	1.22	1.49	**2.55**

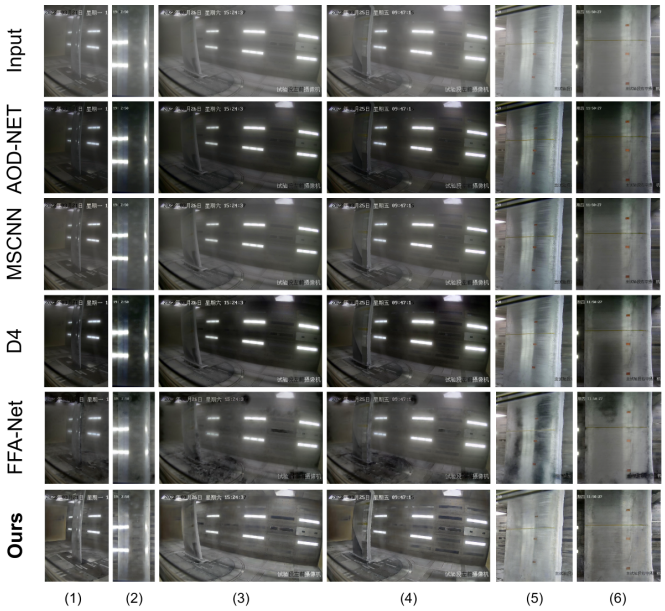

Figure 7. Comparison of six real hazy images in an icing wind tunnel with existing deep learning methods. (1) and (2): LWC=1.31; (3) and (4): LWC=1.0; (5) and (6): LWC=0.5.

5. CONCLUSION

In this paper, we proposed a high-resolution densely connected GAN model for image dehazing in icy meteorological environments. Based on the framework of generative confrontation network, we add sub-pixel convolution for the generator to eliminate the artifacts in the hazy image and obtain high resolution dehazing image. Furthermore the Patch-GAN discriminator can discriminate the image locally to obtain a more accurate haze-free map. Experimental results on icing wind tunnel data demonstrate that our method outperforms the state-of-the-art techniques.

ACKNOWLEDGMENTS

This work was supported by Key Laboratory of Icing and Anti/De-icing of CARDC (Grant No.IADL20210203) and Natural Science Foundation of Sichuan, China under Grant 2023NSFSC0504. All data in this paper are supported by Key Laboratory of Icing and Anti/De-icing of CARDC. Moreover, the authors would like to acknowledge the following people for their assistance: Dr. Quan Zhang, Jiatian Wei, Tianfei Wang, and Shun Wang .

REFERENCES

[1] McCartney, E. J., "Optics of the atmosphere: scattering by molecules and particles," *New York* (1976).

[2] He, K., Sun, J., and Tang, X., "Single image haze removal using dark channel prior," *IEEE transactions on pattern analysis and machine intelligence* **33**(12), 2341–2353 (2010).

[3] Zhu, Q., Mai, J., and Shao, L., "A fast single image haze removal algorithm using color attenuation prior," *IEEE transactions on image processing* **24**(11), 3522–3533 (2015).

[4] Cai, B., Xu, X., Jia, K., Qing, C., and Tao, D., "Dehazenet: An end-to-end system for single image haze removal," *IEEE Transactions on Image Processing* **25**(11), 5187–5198 (2016).

[5] Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., and Yang, M.-H., "Single image dehazing via multi-scale convolutional neural networks," in [*European conference on computer vision*], 154–169, Springer (2016).

[6] Zhang, H. and Patel, V. M., "Densely connected pyramid dehazing network," in [*Proceedings of the IEEE conference on computer vision and pattern recognition*], 3194–3203 (2018).

[7] Guo, T., Li, X., Cherukuri, V., and Monga, V., "Dense scene information estimation network for dehazing," in [*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*], 0–0 (2019).

[8] Ancuti, C. O., Ancuti, C., Sbert, M., and Timofte, R., "Dense haze: A benchmark for image dehazing with dense-haze and haze-free images," *arXiv* (2019).

[9] Jobson, D. J., Rahman, Z.-u., and Woodell, G. A., "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image processing* **6**(7), 965–976 (1997).

[10] Soni, B. and Mathur, P., "An improved image dehazing technique using clahe and guided filter," in [*2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)*], 902–907, IEEE (2020).

[11] Dong, Y., Liu, Y., Zhang, H., Chen, S., and Qiao, Y., "Fd-gan: Generative adversarial networks with fusion-discriminator for single image dehazing," in [*Proceedings of the AAAI Conference on Artificial Intelligence*], **34**(07), 10729–10736 (2020).

[12] Liu, X., Ma, Y., Shi, Z., and Chen, J., "Griddehazenet: Attention-based multi-scale network for image dehazing," in [*Proceedings of the IEEE/CVF international conference on computer vision*], 7314–7323 (2019).

[13] Qin, X., Wang, Z., Bai, Y., Xie, X., and Jia, H., "Ffa-net: Feature fusion attention network for single image dehazing," in [*Proceedings of the AAAI Conference on Artificial Intelligence*], **34**(07), 11908–11915 (2020).

[14] Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q., "Densely connected convolutional networks," in [*Proceedings of the IEEE conference on computer vision and pattern recognition*], 4700–4708 (2017).

[15] Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A. P., Bishop, R., Rueckert, D., and Wang, Z., "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in [*Proceedings of the IEEE conference on computer vision and pattern recognition*], 1874–1883 (2016).

[16] Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A., "Unpaired image-to-image translation using cycle-consistent adversarial networks," in [*Computer Vision (ICCV), 2017 IEEE International Conference on*], (2017).

[17] Hautiere, N., Tarel, J.-P., Aubert, D., and Dumont, E., "Blind contrast enhancement assessment by gradient ratioing at visible edges," *Image Analysis & Stereology* **27**(2), 87–95 (2008).

[18] Galdran, A., "Image dehazing by artificial multiple-exposure image fusion," *Signal Processing* **149**, 135–147 (2018).

[19] Raikwar, S. C. and Tapaswi, S., "Lower bound on transmission using non-linear bounding function in single image dehazing," *IEEE Transactions on Image Processing* **29**, 4832–4847 (2020).

[20] Li, B., Peng, X., Wang, Z., Xu, J., and Feng, D., "Aod-net: All-in-one dehazing network," in [*Proceedings of the IEEE international conference on computer vision*], 4770–4778 (2017).

[21] Yang, Y., Wang, C., Liu, R., Zhang, L., Guo, X., and Tao, D., "Self-augmented unpaired image dehazing via density and depth decomposition," in [*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*], 2037–2046 (2022).