ED-Dehaze Net: Encoder and Decoder Dehaze Network

Hongqi Zhang[1,2*], Yixiong Wei[1,2], Hongqiao Zhou[1,2], Qianhao Wu[1,2]

[1] Department of Engineering & Technology Center, No.38 Research Institute of CETC, Hefei 230088 (China)
[2] Anhui Technical Standard Innovation Base (Intelligent Design and Manufacturing, Intelligence Institute, Civil-military Integration), Hefei 230088 (China)

Received 13 December 2021 | Accepted 22 April 2022 | Published 11 August 2022

Abstract

The presence of haze will significantly reduce the quality of images, such as resulting in lower contrast and blurry details. This paper proposes a novel end-to-end dehazing method, called Encoder and Decoder Dehaze Network (ED-Dehaze Net), which contains a Generator and a Discriminator. In particular, the Generator uses an Encoder-Decoder structure to effectively extract the texture and semantic features of hazy images. Between the Encoder and Decoder we use Multi-Scale Convolution Block (MSCB) to enhance the process of feature extraction. The proposed ED-Dehaze Net is trained by combining Adversarial Loss, Perceptual Loss and Smooth L1 Loss. Quantitative and qualitative experimental results showed that our method can obtain the state-of-the-art dehazing performance.

Keywords

Dehaze, Encoder and Decoder Network, Generative Adversarial Networks, Multi-Scale Convolution Block, Loss Function.

DOI: 10.9781/ijimai.2022.08.008

I. Introduction

IMAGES with clear visibility are required for a variety of computer vision tasks, such as object detection and autonomous driving. However, due to the absorption or reflection of light by floating particles contained in the air, images taken in hazy days often suffer from quality degradation. Fig. 1 shows the differences in visual quality on hazy and haze-free scene. The color of objects in a hazy scene is distorted and the visual perception observed by the human eye is reduced.

In order to overcome the degradation of image quality caused by haze, various priority-based [1]–[5] and learning-based [6]–[14] methods have been proposed. The well-known priority-based algorithm proposed by He et al. [1] assumed that at least one channel in the image has very low pixel values. However, the method cannot effectively deal with areas similar to atmospheric light, which results in the sky area or high-brightness objects cannot be accurately dehazed. In addition, the non local color prior proposed by Berman et al. [3] is suitable for the case where the airlight is lower than the scene brightness [11]. Some recent dehazing algorithms use convolutional neural networks to extract image features, which are further used to predict transmission maps and atmospheric light values. Zhang et al. [9] embed the atmospheric scattering model into the dehazing process and design the end-to-end network with dense connections. By combining the convolutional neural network and the physical model, Zhang's method can jointly learn the transmission map, atmospheric light through a one-stage training.

Inspired by the successful application of the GANs in the field of image generation [15]–[19], we propose ED-Dehaze Net, which contains two parts: a Generator and a Discriminator. The predicted haze-free images are generated by the Generator. The Discriminator is responsible for distinguishing the generated images from the real images. In addition, our Generator consists of three submodules: Encoder, Decoder and Multi-Scale Convolution Block. The Generator and Discriminator are trained simultaneously by combining Adversarial loss, Perceptual loss and Smooth L1 loss. In order to prove the effectiveness of the proposed dehazing method, we conduct sufficient experiments on indoor and outdoor dataset. By using the atmospheric scattering model, we synthesize pairs of hazy and haze-free images on the NYU-Depth v2 [20] indoor dataset. In addition, the commonly used benchmark O-HAZE [21] is also used to verify the performance of ED-Dehaze Net in outdoor dehazing task.

Fig. 1. The effect of haze on the visual quality of the real world scene. From left to right: hazy image, dehazed image produced by ED-Dehaze Net, ground truth haze-free image.

Our main contributions are as follows:

(1) We propose a novel Encoder and Decoder Dehaze Network (ED-Dehaze Net), which can effectively remove the haze in the image.
(2) By using Multi-Scale Convolution Block (MSCB), the feature extraction capability of the dehazing network is improved.
(3) We combine Smooth L1 loss and Perceptual loss to train the Generator. And the experimental results show that the proposed ED-Dehaze Net can obtain the state-of-the-art dehazing performance.

* Corresponding author.
E-mail address: ahu0086@163.com

II. Related Work

In this section, we first review the existing dehazing methods for single image. Then, the research of generative adversarial networks will be briefly summarized.

A. Atmosphere Scattering Model

The Atmospheric scattering model [22]–[24] provides a theoretical basis for the research of dehazing algorithms. Meanwhile, the researchers synthesized hazy images data [9], [10] through the atmospheric scattering model, avoiding the expensiveness of real-world data collection. Its formula is as follows:

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

where x represents the position of pixels, and A means the global atmospheric light in the image. $I(x)$ denotes the haze scene, and $J(x)$ is the haze-free image that the dehazing algorithm expects to obtain. $t(x)$ stands for the medium transmission map, which formula is as follows:

$$t(x) = e^{-\beta d(x)} \tag{2}$$

where β and $d(x)$ represent the atmosphere scattering parameter and the scene depth, respectively.

B. Single Image Dehazing

Single image dehazing is a challenging ill-posed problem. Various prior-based and learning-based methods have been proposed for recovering a clear haze-free image from a single hazy image.

Researchers have proposed a variety of hand-crafted prior-based dehazing methods [1]–[5]. He et al. [1] proposed Dark-Channel Prior (DCP) which can estimate the transmission map and remove the haze effectively. Zhu et al. [2] designed Color Attenuation Prior (CAP) and used a linear model to estimate the scene depth. Then, CAP dehazed a single image by combining the atmospheric scattering model. Fattal et al. [4] proposed a local formation model and use it for recovering the scene transmission. Berman et al. [3] assumed that colors of a haze-free image can be well approximated by hundred of distinct colors. By using the assumption of haze-lines, Berman's method could recover both the distance maps and the haze-free images.

With the development of deep learning, some dehazing algorithms [6]–[14] use convolutional neural networks to learn how to remove haze from hazy images. Li et al. [10] designed All-in-One Dehazing Network (AOD-Net) based on a re-formulated atmospheric scattering model without an intermediate parameter estimation process. On the basis of cycleGAN [25], a cycle-consistent dehazing network [7] is proposed, which does not require paired data. Zhang et al. [9] designed Densely Connected Pyramid Dehazing Network (DCPDN), which can jointly learn the transmission map and atmospheric light. By jointly estimating two parameters, DCPDN can reduce system errors significantly. Dong et al. [14] proposed Multi-Scale Boosted Dehazing Network (MSBDN) based on the U-Net architecture. MSBDN proved that boosting strategy can help image dehazing algorithms have more stable performance.

C. Generative Adversarial Networks (GANs)

Goodfellow et al. [26] proposed the original GAN, which contained a generator (G) and a discriminator (D) for data generation. The objective function is calculated as follows:

$$\begin{aligned} L_{GAN} &= \min_G \max_D V(G, D) \\ &= E_{x \sim p_{\text{data}}(x)}[\log\ D(x)] + E_{z \sim p_z(z)}[\log\,(1 - D(G(z)))] \end{aligned} \tag{3}$$

where z is the noise variable subject to the distribution $p_z(z)$, and x is sampling from the distribution of real data $p_{data}(x)$.

In the recent years, GANs have been successfully applied in the fields of image super-resolution [15]–[17], image synthesis [18], [19], texture synthesis [27], [28], and image inpainting [29], [30]. Most of GANs contain one or more generators and discriminators, and use min-max optimization to simultaneously optimize the generative model G and the discriminative model D [31].

Inspired by the success of these GANs-based methods for generating high-quality images, we designed an adversarial dehazing network with impressive dehazing performance.

III. ED-Dehaze Net

This section introduces the design details of the proposed ED-Dehaze Net. Section A shows the network structure of the Generator and the Discriminator. The formulas of adversarial loss, smooth L1 loss and perceptual loss are given in Section B, C and D, respectively. Finally, Section E describes the overall loss function used for network optimization.

A. Network Structure

The overall structure of the proposed ED-Dehaze Net is shown in Fig. 2. The hazy image passes through the Generator to remove the haze in the scene and generate a dehazed image. Then, the Discriminator distinguishes the dehazed image from the real haze-free image through the process of adversarial training.

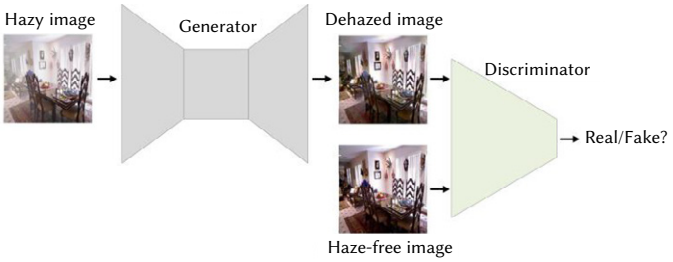

Fig. 2. The overall structure of the ED-Dehaze Net.

The ED-Dehaze Net contains two basic blocks: Convolution Block and Deconvolution Block, as illustrated in Fig. 3. INS stands for Instance Normalization proposed in the paper [32]. RELU is the activation function "Rectified Linear Unit" commonly used in deep neural networks.

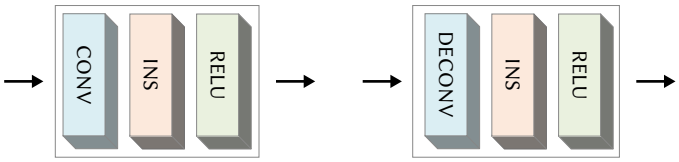

Fig. 3. Convolution Block and Deconvolution Block.

In order to effectively fuse the information of different scales, we design the parallel Multi-Scale Convolution Block (MSCB) as shown in Fig. 4. Each scale of convolution block adopts the Conv Block (shown in Fig. 3) of which 3, 5 and 7 represent convolution kernels of different sizes.

The input feature map first convoluted by Conv-3, Conv-5, Conv-7 to obtain three feature maps with different scale information. Then, the three feature maps are concatenated according to the channel ("Cat" in Fig. 4), and the concatenated 3 map in Fig. 4. Finally, the red feature map and the original input feature map are added ("Add" in Fig. 4). Therefore, the output feature map "OUT" and the input feature map "IN" have exactly the same size.

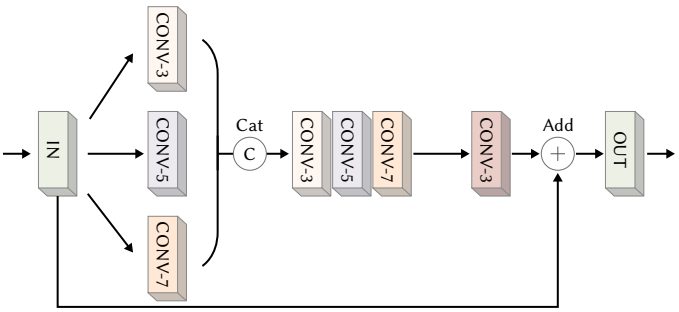

Fig. 4. MSCB: Multi-Scale Conv Block.

The Generator of the ED-Dehaze Net is composed of Encoder, Decoder and multiple MSCBs, as shown in Fig. 5. The dimension of the input image's feature map is reduced during the encoding process to obtain a powerful feature representation. Then, we put the encoded features into MSCBs for enhancing. Finally, the Decoder performs feature decoding to obtain dehazed image of the same size as the original input hazy image.

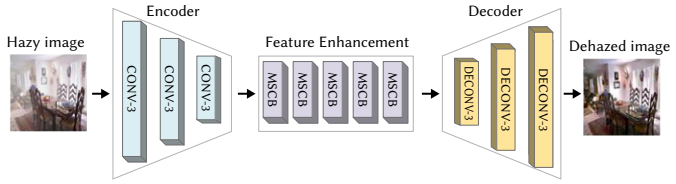

Fig. 5. The Structure of Generator.

The Discriminator shown in Fig. 6 contains multiple Conv-3 blocks, and the stride of all blocks are set to 2. The Discriminator is responsible for distinguishing whether the input images come from the dehazed images of the Generator or the haze-free images of the dataset.

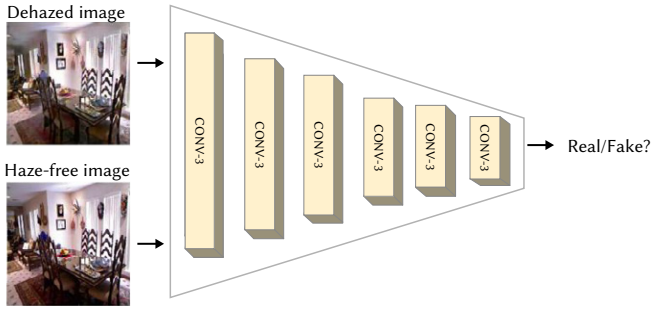

Fig. 6. The Structure of the Discriminator.

B. Adversarial Loss

The Generator and Discriminator are trained simultaneously with adversarial loss. The training purpose of the Generator is to remove the haze in the image to obtain a dehazed image close to the real-world haze-free image. The Discriminator try to judge the input image is a real image or a generated image. The formula for calculating adversarial loss is as follows:

$$L_{\text{adv}} = \min_G \max_D V(G,D) = E_Y \log\, D(Y) + E_X \log\,(1 - D(G(X))) \tag{4}$$

where $X = \{x_1, x_2, ..., x_m\}$ represents the input hazy images with the batch size of m, $\tilde{Y} = \{\tilde{Y}_1, \tilde{Y}_2, ..., \tilde{Y}_m\} = G(X)$ means the haze-free images predicted by the Generator. $Y = \{y_1, y_2, ..., y_m\}$ is the real haze-free labels in the dataset.

C. Smooth L1 Loss

The methods used for image dehazing usually adopt L1-norm or L2-norm as objective function during the training process. Girshick et al. [33] proved that Smooth L1 loss is a robust L1 loss which is less sensitive to outliers than the L2 loss. To improve the robustness of the dehazing network, Smooth L1 loss is adopted to optimize the pixel distance between the dehazed image and the real haze-free image, the formulas are as follows:

$$\text{smooth}_{L1}(\delta_i) = \begin{cases} 0.5(y_i - \tilde{y}_i)^2, & \text{if } |y_i - \tilde{y}_i| < 1 \\ |y_i - \tilde{y}_i| - 0.5, & \text{otherwise} \end{cases} \tag{5}$$

$$L_{smo} = \frac{1}{m}\sum_{i=1}^{m} \text{smooth}_{L1}(\delta_i) \tag{6}$$

where $\delta_i = y_i - \tilde{y}_l$ denotes the distance between the dehazed image and the real haze-free image, and m is the number of images.

D. Perceptual Loss

The commonly used pixel-wise objective functions in image reconstruction tasks optimize the network without considering the human visual perceptual quality of the images. By extracting high-level features from the pre-trained convolutional neural network and calculating the semantic distance between the predicted images and the ground truth images, Johnson et al. [34] proposed Perceptual loss, which has been successfully applied in various computer vision tasks [11], [34]–[36]. It has proved that adding perceptual loss on the basis of pixel-wise loss can effectively improve the performance of image reconstruction tasks. We select the first four pooling layers of pre-trained VGG16 [37] (denoted as $\psi(\bullet)$) for feature extraction. The formula for a single batch is as follows:

$$L_p^k = \frac{1}{m}\sum_{i=1}^{m} \|\psi_k(y_i) - \psi_k(\tilde{y}_i)\|_2 \tag{7}$$

where i represents the i-th input image and k is the k-th pooling layer. And p is the abbreviation of perceptual. After calculating the feature of every single pooling layer, the output of the four pooling layers is added to obtain the Perceptual loss. It can be expressed as following:

$$L_p = \frac{1}{4}\sum_{k=1}^{4} \frac{1}{C_k \times H_k \times W_k} L_p^k \tag{8}$$

where C_k, W_k and H_k represent the number of channels, width and height of the k-th pooling layer, respectively.

E. Overall Loss Function

The overall loss function consists of three parts: Adversarial loss, Perceptual loss, and Smooth L1 loss. As follows:

$$L_{\text{total}} = L_{\text{adv}} + \lambda_{\text{p}} L_{\text{p}} + \lambda_{\text{smo}} L_{\text{smo}} \tag{9}$$

where λ_p and λ_{smo} are the weight of L_p and L_{smo}, respectively.

IV. Experiments Results

Section A gives the parameter settings in the experiment. The descriptions of the synthetic dataset and the real world dataset are in Section B, and the corresponding experimental results are in Section C and D, respectively. To illustrate the effect of Smooth L1 loss and Perceptual loss on the dehazing process, Section E compares the corresponding quantitative metric values by comparing different objective functions. Finally, Section F discusses the improvement of MSCBs on the performance of ED-Dehaze Net.

A. Experiment Setting and Evaluation

All the training images are resized to 256×256 before feeding into the network. The proposed model has been trained on the training dataset and finally the optimal parameters are obtained. The initial learning rate of the Generator and Discriminator are 0.0001 and 0.0004, respectively. After every 2000 iterations, the learning rate is updated to lr = lr × 0.9. The values of λp and λsmo are set to 2 and 10, respectively. Wet use a single NVIDIA GeForce GTX 1080 Ti and the batch size is 4. Adam Optimizer is used to update the gradients during the training process.

We choose Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) for evaluation, which are commonly used in the research of image dehazing. Higher values of PSNR and SSIM represent higher image quality. To demonstrate the improvement of the dehazing performance for the proposed ED-Dehaze Net, we compare our method with other state-of-the-art dehazing methods: the Dark-Channel Prior (DCP) proposed by He et al. [1], the Color Attenuation Prior (CAP) proposed by Zhu et al. [2], the cycle-consistency dehazing (cyc-D) network designed by Engin et al. [7], and the All-in-One Dehazing Network (AOD-Net) designed by Li et al. [10].

B. Synthesis Dataset and Real World Dataset

The collection of paired hazy and haze-free images is very time-consuming and expensive. Ancuti et al. [38] synthesized the indoor hazy images dataset based on NYU Depth v2 [20] by the atmospheric scattering model. Here, we use the pipeline proposed by Aucuti for high quality hazy images synthesis. The training set and testing set contain 1149 and 300 pairs of images, respectively. We randomly select the global atmosphere light value from (0.7, 1) for both of them. The atmosphere scattering parameter (denoted as β in Eq. (2)) of the training dataset is set to β ∈ (1.2, 2.1) to simulate hazy images of different densities. By randomly flipping horizontally and vertically, the training set after data augmentation contains 1149 × 3 pairs of images. In order to test the dehazing performance of networks with different density ranges and to ensure the data distribution of the testing set is same as training set, we set the β ∈ (1.2, 1.5), (1.5, 1.8), (1.8, 2.1), and (1.2, 2.1) to get a testing set with 4 density ranges for a total of 300 × 4 pairs of images.

O-HAZE [21] contains 45 pairs of hazy and haze-free images taken in the real world, which the haze is generated by professional machines. We randomly selected 35 pairs as the training set, and the remaining 10 pairs as the testing set.

C. Results on Synthesis Dataset

Table I shows the PSNR/SSIM values obtained on the synthetic dataset by the proposed ED-Dehaze Net and other dehazing algorithms. ED-Dehaze Net outperforms the state-of-the-art methods on both PSNR and SSIM. This indicates that the dehazing images obtained by our method are of higher quality.

TABLE I. PSNR and SSIM Values on Synthesis Dataset

beta	(1.2, 1.5)	(1.5, 1.8)	(1.8, 2.1)	(1.2, 2.1)
DCP	16.*314*/[#]0.761	16.212/0.776	*15.*935/**0.783**	*16.*346/0.771
CAP	17.422/0.768	17.062/0.723	*16.*654/0.702	*17.*175/0.753
cyc-D	17.553/0.657	18.074/0.694	18.263/0.*725*	17.934/0.766
AOD	18.736/0.782	18.445/**0.788**	17.*963*/0.779	18.566/0.761
ours	**20.940**/**0.799**	**20.332**/0.784	**19.627**/0.761	**20.264**/**0.781**

[#] The symbol "/" stands for the separation of PSNR and SSIM values.

We randomly select some images from the entire testing set (four ranges of β) to show the visual results. It can be clearly seen in Fig. 7 that the images recovered by the proposed ED-Dehaze Net are closest to the ground truth haze-free images. This proves that the Encoder-Decoder of the Generator can effectively reconstructs the feature information of the image, and removes the haze contained in the scene in an end-to-end manner. The DCP results in a darker color after dehazing, which cannot accurately recover the brightness of the scene. The reason is that DCP relies on priority assumptions and lacks adaptability to different data. The CAP and AOD-Net cannot completely remove the haze in the hazy images. There is still a small amount of haze in the "haze-free" images generated by them, which causes the details and edges of objects in the scenes to be blurred. In addition, the output images of CAP tends to be low-brightness, because its priori assumption of color is not always accurate. The dehazing images generated by cyc-D have obvious color distortion, and it cannot restore the texture information of the image completely.

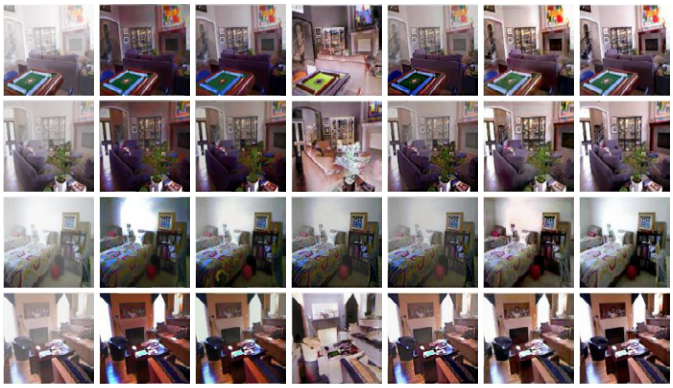

Fig. 7. Visual results on synthesis dataset.

D. Results on O-HAZE

Table II shows the PSNR/SSIM values of the ED-Dehaze Net and other algorithms on the O-HAZE dataset. The quantitative quality of the dehazed images obtained by our method is significantly higher on outdoor scenes. This proves that ED-Dehaze Net can reduce the proportion of noise and restore the structure information of the image.

TABLE II. PSNR and SSIM Values on O-HAZE

Methods	DCP	CAP	AOD	cyc-D	ours
PSNR	16.160	17.078	17.376	19.626	**19.925**
SSIM	0.772	0.792	0.786	0.677	**0.795**

Fig. 8 shows the visual results of various dehazing methods on the O-HAZE dataset. Both DCP and CAP have a phenomenon of color shift and there is still a small amount of haze remaining in the dehazed images, resulting in blurred details. Although AOD-Net uses a convolutional neural network for feature extraction, it does not use multi-scale spatial information like ED-Dehaze Net. So, the AOD-Net can remove the haze in the images, but cannot accurately reconstruct the texture and details. The dehazing images of the cyc-D network has over-smooth, because its unsupervised strategy cannot accurately estimate the density of haze in the image, resulting in excessive dehazing. Meanwhile, Table II shows that the SSIM value of the cyc-D is low, which proves that its excessive dehazing leads to incomplete restoration of the images' structure. Our method can obtain high-quality haze-free images, ensuring that the edge information contained in the images will not be lost.

Fig. 8. Visual results on O-HAZE.

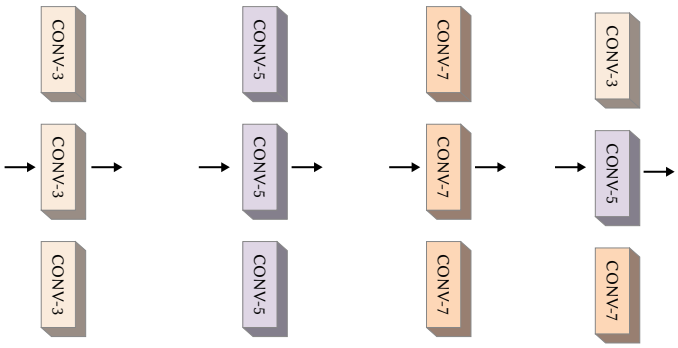

Fig. 9. Comparison with different scales, from left to right: 3 × 3, 5 × 5, 7 × 7 and our multi-scale 3 + 5 + 7.

E. Ablation Study on Loss Function

During the training of the Generator, Smooth L1 loss and Perceptual loss are used to optimize the network. To prove that with the help of Perceptual loss, the network with Smooth L1 loss can generate dehazed images with higher quality. We compare two strategies when choosing loss functions: (a) using Smooth L1 loss alone, and (b) using Smooth L1 loss and Perceptual loss simultaneously. We have not chosen Perceptual loss as the loss function of the Generator alone, because the purpose of Perceptual loss is to measure the distance of features rather than the distance of image pixels. Therefore, the role of Perceptual loss should be an auxiliary of the pixel-wise Smooth L1 loss. Table III lists the ablation experiment results of Smooth L1 loss and Perceptual loss.

TABLE III. Ablation Experiment Results of Smooth L1 Loss and Perceptual Loss

beta	*(1.2, 1.5)*	*(1.5, 1.8)*	*(1.8, 2.1)*	*(1.2, 2.1)*
Smo	19.667/#0.774	19.106/0.724	19.126/0.756	19.*644*/0.760
L1+ Per	19.821/0.782	19.647/0.781	18.966/**0.772**	20.032/0.774
L2+ Per	19.731/0.748	**20.892**/0.768	19.614/0.758	20.201/0.769
Smo+Per	**20.940**/**0.799**	20.332/**0.784**	**19.627**/0.761	**20.264**/**0.781**

The symbol "/" stands for the separation of PSNR and SSIM values.

The Smo+Per get a higher PSNR/SSIM values than Smo, which means Perceptual loss can help the network generate high-quality dehazed images. In addition, the PSNR/SSIM values obtained by L1 + Per and L2 + Per are lower than Smo + Per. This proves that Smooth L1 loss can optimize the Generator better.

F. Ablation Study on Multi-Scale Convolution Block

The Generator contains multiple Multi-Scale Convolution Blocks (MSCBs) to enhance the process of the feature extraction, so as to obtain better dehazing performance. In order to prove the effectiveness of the multi-scale strategy, we compared the PSNR/SSIM values obtained by a single scale (3 × 3, 5 × 5, 7 × 7) and multiple scales (3 + 5 + 7) as shown in Fig. 9. In order to ensure the fairness of the experiment, all other parameter settings are the same.

The results in Table IV prove that our multi-scale strategy can extract features more effectively. Therefore, ED-Dehaze Net with MSCBs can obtain higher quality dehazed images. For the three single-scale cases (3 × 3, 5 × 5, 7 × 7), the PSNR/SSIM values are very close. This further illustrates that the feature fusion method of multi-scale convolution can enhance the flow of spatial information.

TABLE IV. Ablation Experiment Results of MSCBs

beta	*(1.2, 1.5)*	*(1.5, 1.8)*	*(1.8, 2.1)*	*(1.2, 2.1)*
3 × 3	**21.033**/0.781	19.634/0.792	19.231/0.755	20.132/0.779
5 × 5	19.872/0.796	19.897/0.766	18.998/0.755	19.245/0.773
7 × 7	19.660/0.790	20.226/0.793	19.046/0.758	19.995/0.760
ours	20.940/**0.799**	**20.332**/0.784	**19.627**/**0.761**	**20.264**/**0.781**

Meanwhile, we conduct experiments on the case where the Generator contains only Encoder-Decoder and no MSCBs. The results in Table V show that PSNR and SSIM are very low. The main reason is that the network's capacity and complexity are not sufficient when without MSCBs.

TABLE V. Results With MSCBs and Without MSCBs

beta	*(1.2, 1.5)*	*(1.5, 1.8)*	*(1.8, 2.1)*	*(1.2, 2.1)*
noMSCBs	18.556/0.749	18.673/0.715	17.143/0.729	17.011/0.712
ours	20.940/0.799	20.332/0.784	19.627/0.761	**20.264**/**0.781**

V. Conclusion

This paper proposed an end-to-end dehazing algorithm based on deep learning, called ED-Dehaze Net. The Generator effectively extracted the spatial and texture information of the hazy images by the Encoder-Decoder structure. In order to ensure that the edges and details of the dehazed image are clearly reconstructed, we used Smooth L1 Loss and Perceptual Loss to train the Generator simultaneously. Experiments on synthetic and real world dataset proved that the proposed algorithm can effectively remove the haze in the hazy images.

Acknowledgment

We are very grateful to Dr Jun Zhang and Dr Peng Chen, professors in Anhui University, for their invaluable helps to develop the proposed method and prepare the original manuscript.

References

[1] K. He, J. Sun, X. Tang, "Single image haze removal using dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 2341-2353, 2010.

[2] Q. Zhu, J. Mai, L. Shao, "A fast single image haze removal algorithm using color attenuation prior," *IEEE Transactions on Image Processing*, vol. 24, pp. 3522 -3533, 2015.

[3] D. Berman, S. Avidan, S. Avidan, "Non-local image dehazing," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 1674 -1682.
[4] R. Fattal, "Dehazing using color-lines," *ACM Transactions on Graphics*, vol. 34, pp. 13, 2014.
[5] G. Meng, Y, Wang, J. Duan, S. Xiang, C. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," in *IEEE International Conference on Computer Vision*, 2013, pp.617-624.
[6] B. Cai, X. Xu, K. Jia, C. Qing, D. Tao, "Dehazenet: An end-to-end system for single image haze removal," *IEEE Transactions on Image Processing*, vol. 25, pp. 5187 -5198, 2016.
[7] D. Engin, A. Genc, H. Kemal Ekenel, "Cycle-dehaze: Enhanced cyclegan for single image dehazing," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 825 -833.
[8] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, M.H. Yang, "Single image dehazing via multi-scale convolutional neural networks," in *European Conference on Computer Vision*, Springer, 2016, pp. 154 -169.
[9] H. Zhang, V.M. Patel, "Densely connected pyramid dehazing network," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 3194 -3203.
[10] B. Li, X. Peng, Z. Wang, J. Xu, D. Feng, "An all-in-one network for dehazing and beyond," 2017, *arXiv*: 1707.06543.
[11] X. Yang, Z. Xu, J. Luo, "Towards perceptual image dehazing by physics-based disentanglement and adversarial training," in *AAAI conference on artificial intelligence*, 2018, pp. 7485 -7492.
[12] X. Qin, Z. Wang, Y. Bai, X. Xie, H. Jia, "Ffa-net: Feature fusion attention network for single image dehazing," 2019 *arXiv*: Computer Vision and Pattern Recognition.
[13] S. Yin, Y. Wang, Y.H. Yang, "A novel residual dense pyramid network for image dehazing," *Entropy*, vol. 21, pp. 1123, 2019.
[14] H. Dong, J. Pan, L. Xiang, Z. Hu, X. Zhang, F. Wang, M.H. Yang, "Multi-scale boosted dehazing network with dense feature fusion," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (CVPR), 2020, pp. 2154-2164, doi: 10.1109/CVPR42600.2020.00223.
[15] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, C. Change Loy, "Esrgan: Enhanced super-resolution generative adversarial networks," in *European Conference on Computer Vision*, 2018, pp. 63 -79.
[16] Y. Yuan, S. Liu, J. Zhang, Y. Zhang, C. Dong, L. Lin, "Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 701-710.
[17] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al. "Photo-realistic single image superresolution using a generative adversarial network," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4681 -4690.
[18] R. Huang, S. Zhang, T. Li, R. He, "Beyond face rotation: Global and local perception gan for photorealistic and identity preserving frontal view synthesis," in *International Conference on Computer Vision*, 2017, pp. 2439 -2448.
[19] L. Ma, X. Jia, Q. Sun, B. Schiele, T. Tuytelaars, L. Van Gool, "Pose guided person image generation," in *Neural Information Processing Systems*, 2017, pp. 406-416.
[20] N. Silberman, D. Hoiem, P. Kohli, R. Fergus, "Indoor segmentation and support inference from rgbd images," in *European Conference on Computer Vision*, Springer, 2012, pp. 746 -760.
[21] C.O. Ancuti, C. Ancuti, R. Timofte, C. De Vleeschouwer, "O-haze: a dehazing benchmark with real hazy and haze-free outdoor images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 754-762.
[22] E.J. McCartney, "Optics of the atmosphere: scattering by molecules and particles," New York, John Wiley and Sons, Inc., 1976.
[23] S.K. Nayar, S.G. Narasimhan, "Vision in bad weather," in *IEEE International Conference on Computer Vision*, 1999, pp. 820-827.
[24] S.G. Narasimhan, S.K. Nayar, "Contrast restoration of weather degraded images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 6, pp. 713 -724, 2003.
[25] J.Y. Zhu, T. Park, P. Isola, A.A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *International Conference on Computer Vision*, 2017, pp. 2223 -2232.
[26] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, "Generative adversarial networks," *Advances in Neural Information Processing Systems*, vol. 3, pp. 2672 -2680, 2014.
[27] C. Li, M. Wand, "Precomputed real-time texture synthesis with markovian generative adversarial networks," in *European Conference on Computer Vision*, 2016, pp. 702 -716.
[28] N. Jetchev, U. Bergmann, R. Vollgraf, "Texture synthesis with spatial generative adversarial networks," *arXiv*: 1611.08207, 2016.
[29] R.A. Yeh, C. Chen, T. Yian Lim, A.G. Schwing, M. Hasegawa-Johnson, M.N. Do, "Semantic image inpainting with deep generative models," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 5485 -5493.
[30] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, T.S. Huang, "Generative image inpainting with contextual attention," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 5505 -5514.
[31] J. Gui, Z. Sun, Y. Wen, D. Tao, J. Ye, "A review on generative adversarial networks: Algorithms, theory, and applications," *arXiv*: Learning, 2020.
[32] D. Ulyanov, A. Vedaldi, V. Lempitsky, "Instance normalization: The missing ingredient for fast stylization," *arXiv*: Computer Vision and Pattern Recognition, 2016, pp. 1607.08022.
[33] R. Girshick, "Fast R-CNN," in *The IEEE International Conference on Computer Vision* (ICCV), 2015, pp. 1440-1448, doi: 10.1109/ICCV.2015.169.
[34] J. Johnson, A. Alahi, L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *European Conference on Computer Vision*, Springer, 2016, pp. 694 -711.
[35] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M.K. Kalra, Y. Zhang, L. Sun, G. Wang, "Low-dose ct image denoising using a generative adversarial network with wasserstein distance and perceptual loss," *IEEE Transactions on Medical Imaging*, vol. 37, pp. 1348 -1357, 2018.
[36] G. Yang, Y. Cao, X. Xing, M. Wei, "Perceptual Loss Based Super-Resolution Reconstruction from Single Magnetic Resonance Imaging," In: X. Sun, Z. Pan, E. Bertino, (eds) Artificial Intelligence and Security. ICAIS 2019. Lecture Notes in Computer Science, vol. 11632, p. 411–424.
[37] K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv*: 1409.1556, 2014.
[38] C. Ancuti, C.O. Ancuti, C. De Vleeschouwer, "D-hazy: A dataset to evaluate quantitatively dehazing algorithms," in *IEEE International Conference on Image Processing*, 2016, pp. 2226-2230.

HongQi Zhang

He is currently chief expert of China Electric Power Group, researcher professor of No. 38 Research Institute of CETC, ISO registered expert, and expert of National Defense Science and Industry Bureau, National Standard Commission, the Ministry of Science and Technology as well as the Ministry of Industry and Information Technology. He leads more than 80 items for the formulation of International, national and industry standards, including 4 international standards, 32 national standards, electronics and machinery industry standards more than 50 items. He won more than 20 provincial-level science and technology awards, including 4 China Standard Innovation Contribution Awards. In 2020, he won the China Standard Innovation Contribution Award, Outstanding Contribution Award; in 2018, he won the China Standard Innovation Contribution Award, with ranking the first order.

Yixiong Wei

He is the senior engineer of No. 38 Research Institute of CETC, and the master supervisor of Chinese Academy of Electrical Sciences. He is long-term engaged in digital cutting-edge technology research, technology breakthrough, application transformation and other works. He presided over the "air 500 Dam Components Assembly Line Intelligent Information Iot" and other projects, as well as the chief designer and other core members to participate in the bureau of Science and technology and other 7 projects. He has obtained 18 patent authorizations, including 1 International Patent Authorization and 6 National Patent Authorizations. He participates in the preparation of 29 standards at various levels, including 1 national standard as the chief editor, 3 industry standards. He has won the first prize of Science and technology progress of Anhui Province, the excellent patent award of China, the first prize of China Quality Association, and the First Prize of science and technology of China Electronic Science and Technology Group, etc.

Zhou Hongqiao

He received the B.S. degree and M.S. degree in the mechanical design manufacture and automation from Wuhan University, Wuhan, China, in 2002 and in 2005, and the Ph.D. degree in mechanical manufacture and automation from Huazhong University of Science and Technology, Wuhan, China, in 2010. He is currently a researcher with the Institute of Electronics Technology, The 38th Research Institute of China Electronics Technology Group Corporation, Hefei, China. He is the author of 8 articles, and 12 national invention patents. He participated in the designation of 2 national standards and 51 military electronic industry standards. His research interests include the designation of electronic industry standards, digital design and manufacturing, enterprise information theory and system integration researches. Dr. Zhou was a recipient of the Second Prize of Anhui Province Science and Technology Progress Award in 2016, and the China Patent Excellence Award in 2017, and the First Prize of China Standard Innovation Contribution Award in 2018.

Wu Qianhao

He was born in Anqing, Anhui, China in 1995. He received the bachelor's degree in the mechanical design manufacture and automation from Anhui University, Hefei, China, in 2017 and the master's degree in transportation engineering from Hefei University of Technology, Hefei, China, in 2020. He is currently a digital design engineer with the Institute of Electronics Technology, The 38th Research Institute of China Electronics Technology Group Corporation, Hefei, China. His research interests include digital twin technology, digital technology and intelligent manufacturing.