

Received October 27, 2021, accepted November 15, 2021, date of publication November 23, 2021, date of current version December 6, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3130468

A Progressive Single-Image Dehazing Network With Feedback Mechanism

TISONG LIANG¹, ZHIWEI LI¹, YUANHONG REN², QI MAO¹, AND WUNENG ZHOU²

¹College of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China

²College of Information Science and Technology, Donghua University, Shanghai 200051, China

Corresponding author: Zhiwei Li (zhiwei.li@sues.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61705127, in part by the Shanghai Natural Science Foundation under Grant 19ZR1421600, and in part by the Degree Construction Project of Detection Technology and Automation Devices, Shanghai University of Engineering Science, under Grant 19XXK003.

ABSTRACT In the past decade, deep learning methods, especially convolutional neural networks, have received much attention in applications of single-image dehazing. However, the haze in hazy images cannot be distinctly separated because it is complicatedly mixed with the background components. If we roughly remove the haze, the background tone of global atmospheric light may also be destroyed. To resolve the above problem and reconstruct clearer and higher-quality dehazing images, we introduced our progressive feedback network (PFBN) in recurrent structure ties with a feedback mechanism. The feedback mechanism is implemented by stacking feedback blocks that contain feedback connections among iterations. At the input layer of each feedback block, its hidden state in the last iteration is delivered by a feedback connection to the present block as part of the input. The last hidden state, also referred to as high-level information, is fused with low-level information output by the previous block to generate effective feature representation. Moreover, we proposed an enhancement self-ensemble strategy to decrease the random error of the network to reconstruct clearer dehazing images. Finally, we designed a series of extensive experiments to verify the outstanding performance of our method.

INDEX TERMS Image dehazing, deep learning, feedback mechanism, convolutional neural network.

I. INTRODUCTION

One of the most common types of weather is haze, which generates complicated noise. Therefore, it brings about great challenges if hazy images need to be used in upper applications. Removing haze is necessary for many outdoor applications, such as objective recognition and video surveillance [1]. However, because the amount of useful information in a single hazy image is insufficient, dehazing algorithms are always considered ill-posed tasks [2]. Today, this topic is one of the most attention-grabbing academic explorations in image reconstruction and artificial intelligence [3].

Many algorithms focusing on single-image dehazing have been proposed in recent years. These algorithms can be roughly classified into two categories: model-based approaches and data-driven approaches, that is, deep learning-based methods. According to previous studies [4]–[6], the classical model for generating a hazy image, I ,

is as follows:

$$I(x) = T(x)J(x) - (1 - T(x))A \quad (1)$$

where J refers to haze-free scene brilliance; I refers to the observed hazy image; A denotes the global light of the atmosphere, indicating the strength of the environmental light; T refers to the transmission map; and x is the pixel location.

Most methods use this model and design two networks to obtain A and T ; however, this model is computationally intensive, and the proposed networks are always difficult to train and have strict high PC requirements.

According to [2], a simpler but more effective model is employed in this paper:

$$\begin{aligned} J(x) &= K(x)I(x) - K(x) + b \\ K(x) &= \frac{\frac{1}{T(x)}(I(x) - A) + (A - b)}{I(x) - 1} \end{aligned} \quad (2)$$

In the equation above, $K(x)$ is a new unified variable contains $\frac{1}{T(x)}$ and A . Variable b refers a constant bias and default value is set as 1 as [2] did.

The associate editor coordinating the review of this manuscript and approving it for publication was Sudhakar Radhakrishnan.

Therefore, we only need to estimate one unified parameter matrix $K(x)$ through the network rather than two in Eqn.1. It is more effective and computationally efficient.

Currently, there have been many upper applications in the field of computer vision or image enhancement, and all these applications, such as object detection and autopilot, require a clean and high-quality dataset for their implementation. Therefore, the performance of upper applications depends to a large extent on whether the previous data processing process was successful. Therefore, it is vital to enhance the captured visual data, such as through the use of dehazing and deraining, to pursue cleaner data, which can then be utilized in upper applications.

II. RELATED WORK

A. DATA-DRIVEN SINGLE-IMAGE DEHAZING METHODS

To date, in the field of computer vision and image enhancement, most of the designed networks have been end-to-end networks: introducing a specific network to reconstruct clean images immediately from the original hazy images. There exists a wide consensus that the key point of reconstructing clean images is to obtain a precise medium transmission map. [7], [8].

The recurrent structure is also employed in image dehazing. These methods are aimed at removing haze accumulated in images both iteratively and progressively. Jiang *et al.* [9] proposed a lightly recurrent network for video dehazing, which makes it possible to train such a large dataset in a limited amount of time. Tan *et al.* [10] introduced a novel method in which a random Markov field was employed to obtain transmission map T .

B. FEEDBACK MECHANISM

The feedback mechanism is a simple but effective approach to enhancing the performance of networks in recurrent structures. The core of this mechanism is delivering high-level information to low-level information through a feedback connection to integrate a power feature representation. The output of the network and reconstruction of clean images are then iteratively clearer.

Many previous studies have successfully applied feedback mechanisms in various missions. The author of [11] first introduced a feedback mechanism and successfully adopted it for image classification and designed some helpful strategies that obtained outstanding performance. Later, [12] utilized a feedback strategy for superresolution and replaced the LSTM [13] block with their own feedback block to achieve a better result. Then, Shama *et al.* [14] found a helpful improvement by combining a feedback approach with generative adversarial networks, the feedback loop is employed to transmit the discriminatory spatial information to the generator. Recently, a feedback strategy has been employed in action detection with graph neural networks to enhance the representation ability of networks [15]. [16] proposed a novel feedback network for single image deraining, a residual dense block

(RDB) [17] is employed as feedback block, RDB is effective for image deraining but not helpful for image dehazing according to our experiments.

In this paper, we introduce a unique feedback block to make it apposite with a feedback mechanism. After several feedback blocks are stacked, the designed deep neural network can extract more precise information and features from the original hazy images; then, strong representation can be helpful for reconstructing the dehazing images. Due to the ability to correct the previous state, haze can be eliminated gradually, while the background iteratively becomes clearer.

III. PROPOSED METHODS

Two components are required in a feedback mechanism. 1) iterativeness. 2) rerouting the output back into the system to amend the input in each loop. To meet this demand, the recurrent structure is employed and feedback loops are adopted in each feedback block. As Fig.1, the whole network is folded by T sub-networks, where T refers to the settled parameter that represents the number of iterations. Besides, the parameters of each sub-networks are globally shared across iterations. What's more, each sub-networks will be forced to estimate $K(x)$ and reconstruct a dehazing image for sub-loss calculating.

A. NETWORK STRUCTURE

- 1) Input Layer: The input layer f_{in}^t in the t^{th} iteration receives the original hazy image I and the dehazing image output by the previous iteration J^{t-1} .

$$F_{in}^t = f_{in}^t([J^{t-1}|I]) \quad (3)$$

where $[\cdot|\cdot]$ refers the concatenation operation and F_{in}^t is the feature map produced by the input layer. The input layer essentially a fusion layer that merge the high-level information of its previous iteration and the low-level original information. The input layer consists of a convolutional operation $Conv(6, m)$, where m is the number of channels of F_{in}^t , we set $m = 64$ in this paper.

- 2) Feedback Backbone: The feedback backbone $f_{backbone}^t$ is a combination of several stacked feedback blocks $\{FB_1, FB_2, \dots, FB_B\}$ (The detailed description of feedback block is in III-B and the exact number of feedback blocks will be explored in IV-A). The feedback backbone receives the F_{in}^t output by the input layer.

$$F_{backbone}^t = f_{backbone}^t(F_{in}^t) \quad (4)$$

where $F_{backbone}^t$ refers the hidden states produced by feedback backbone in the t^{th} iteration. It is noticeable that each feedback block in the internal of feedback backbone will receive extra information, more description can be found in III-B.

- 3) Output Layer: In the output layer f_{out}^t , feature map $F_{backbone}^t$ is used to calculate the unified parameter matrix $K^t(x)$.

$$K^t(x) = f_{out}^t(F_{backbone}^t) \quad (5)$$

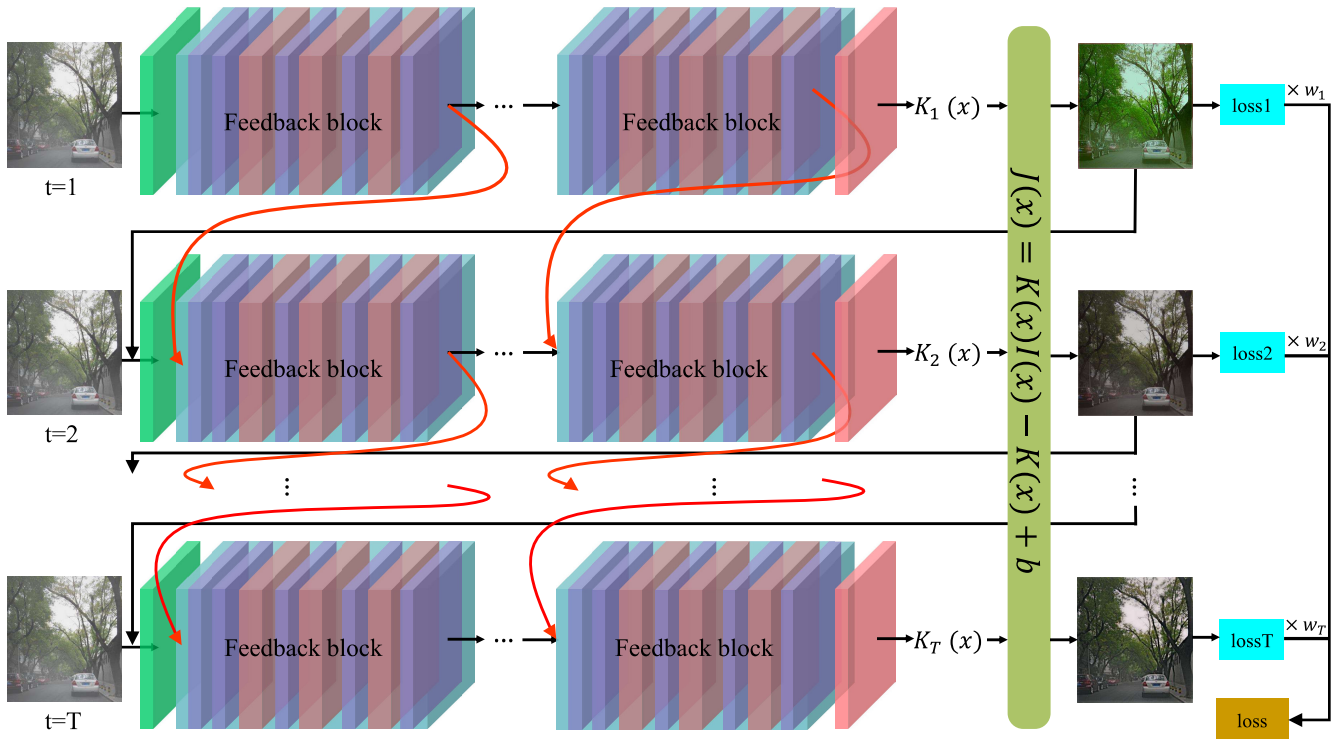


FIGURE 1. Illustration of our proposed method with T iterations. The green block on the left refers the input layer. The pink block on the right refers the output layer. The red lines denotes the feedback connection designed for information transmission. The black lines from dehazy images to input layer indicate that the input layer will receive the reconstruction dehazing image of last iteration. It is notable that the parameters at each iteration are shared globally. The network is forced to restore the background at each iteration and T sub-losses are calculated to obtain the final loss.

The output layer contains a convolutional operation $Conv(m, 3)$ to transform the m -channels $F_{backbone}^t$ into 3-channels $K^t(x)$.

- 4) Recurrent Loss: In each iteration, after obtaining the estimated $K^t(x)$, sub-network will be enforced to generate a hazy-free scene and calculate the sub-loss.

$$loss = \frac{\sum_{i=1}^T w_i loss_i}{\sum_{i=1}^T w_i} \quad (6)$$

where $loss$ is the final loss of the whole network, $loss_i$ is the sub-loss of the i^{th} sub-network, and w_i is the weight corresponding to each sub-loss. In this paper, we set $w_i = 1, i = 1, 2, \dots, T$ as [11] did, and T is the number of iteration of proposed network (The exact value of T will be discussed in IV-A).

The feedback mechanism requires rerouting the output back into the system in each iteration [11]. In that case, the main difference between feedback network and recurrent feedforward network is recurrent loss. If simply adopting the loss of the last iteration as final loss will lead to a feedforward network rather than feedback network. Therefore, every sub-loss of each iteration must be considered in our proposed network.

B. FEEDBACK BLOCK

A feedback block is composed of a fusion layer and a K -estimation module proposed by [2]. The fusion layer of the

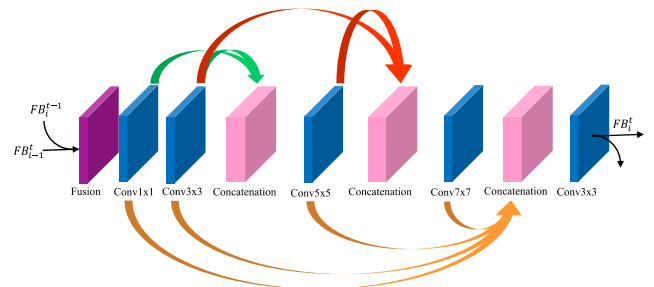


FIGURE 2. A closer look of feedback block.

i^{th} feedback block in the t^{th} iteration will receive the output of the $i - 1^{th}$ feedback block in the t^{th} iteration and the output of the i^{th} feedback block in the $t - 1^{th}$ iteration, denotes as FB_{i-1}^{t-1} and FB_i^{t-1} respectively. The fusion layer is essential a average operator, it will deliver the average feature map of FB_{i-1}^{t-1} and FB_i^{t-1} to K -estimation module.

As shown in Fig.2. The K -estimation contains 5 convolutional operations and 3 concatenation operations. The first concatenation layer will concatenate the hidden states produced by $Conv 1 \times 1$ and $Conv 3 \times 3$. Similarly, the second concatenation layer is responsible for processing those of $Conv 3 \times 3$ and $Conv 5 \times 5$. The final concatenation layer will receive all hidden states produced by previous 4 convolutional operations and deliver the concatenated feature map to the final convolutional layer.

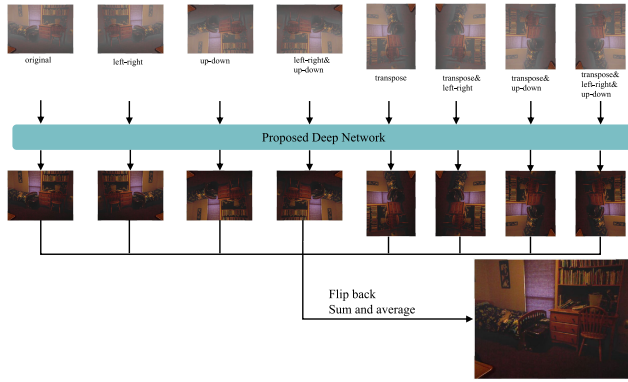


FIGURE 3. The illustration of self-ensemble strategy.

C. SELF-ENSEMBLE

The self-ensemble method is an enhancement method to decrease the variance in the model [18]. By horizontally, vertically, or diagonally flipping the outputs of augmented images, we can obtain eight ensemble hazing images $[I_{B_1}, I_{B_2}, \dots, I_{B_8}]$ as shown in Fig.3. With the inspiration of great success achieved by this strategy in field of super resolution, we hope to adopt this strategy in single image dehazing.

After feeding the network these eight hazing images, we can obtain eight dehazing images $[O_{B_1}, O_{B_2}, \dots, O_{B_8}]$, flip them back and average them to obtain the final output of the network.

$$O = \frac{\sum_{i=1}^8 O_{B_i}}{8} \tag{7}$$

IV. EXPERIMENTS RESULTS

In this section, we conduct a series of ablation experiments to determine the key components of our proposed network (PFBN). Then, we quantitatively and qualitatively evaluate our network on several common benchmark datasets with state-of-the-art methods.

In the details of our proposed method, the number of channels m is set to 64, and the kernel size of convolutional operations is 3. All experiments are implemented with the PyTorch framework [19]. The training processes are conducted on a PC with a Linux system equipped with NVIDIA 2080Ti GPUs and Intel(R) Core(TM) i7-11800H CPU with 16 GB RAM. In the experiments, ADAM optimization [20] is adopted to train the models with an initial learning rate of 1e-5.

A. ABLATION STUDIES

All ablation studies are carried out on a benchmark test dataset called SOTS [12], which contains 1,000 pairs of haze images and cleaning images, the evaluation metrics of which are PSNR and SSIM [21].

1) STUDY OF ITERATIONS

In this subsection, we discuss the influence of the number of iterations (denoted as T) while fixing the number of blocks

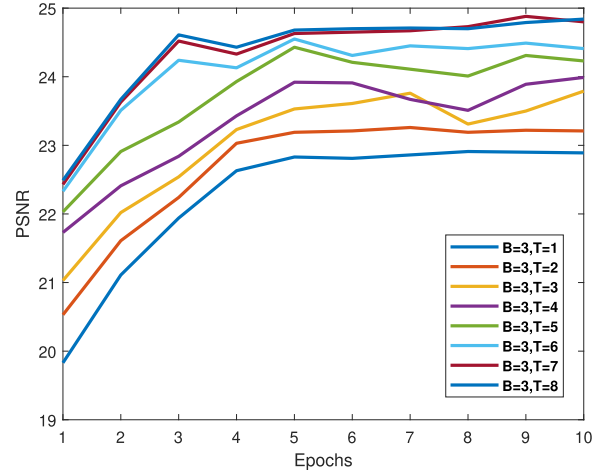


FIGURE 4. PSNR values corresponding to each iteration.

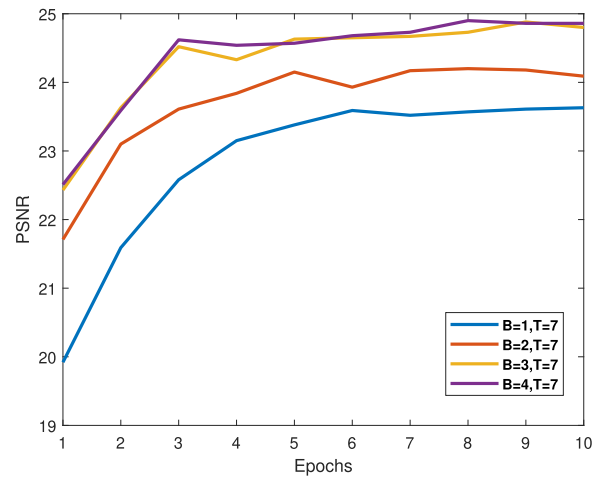


FIGURE 5. PSNR values corresponding to each number of feedback blocks.

as 3. Fig. 4 indicates the PSNR values of the proposed FBN, and $T = 1, 2, 3, \dots, 7, 8$. It is notable that when $T > 1$ (with feedback connection), the metric of reconstruction images is significantly higher than $T = 1$. This means that our proposed PFBN indeed benefits from the feedback strategy. Moreover, when T increases, the performance continues to rise. In addition, we note that when $T > 7$, the performance almost converges to a constant. Although there is still a slight improvement in performance when $T > 7$, we comprehensively consider the GPU memory and running time, and we set $T = 7$ as final number of iteration in proposed network structure.

2) STUDY OF THE NUMBER OF FEEDBACK BLOCK

Similar with the study of iterations, we explore the number of feedback blocks (denoted as B) by settle $B = 1, 2, 3, 4$ while fixing the $T = 7$. Fig.5 shows that $B = 3$ is the most suitable value. So, the final structure of our proposed PFBN is designed as $B = 3, T = 7$.

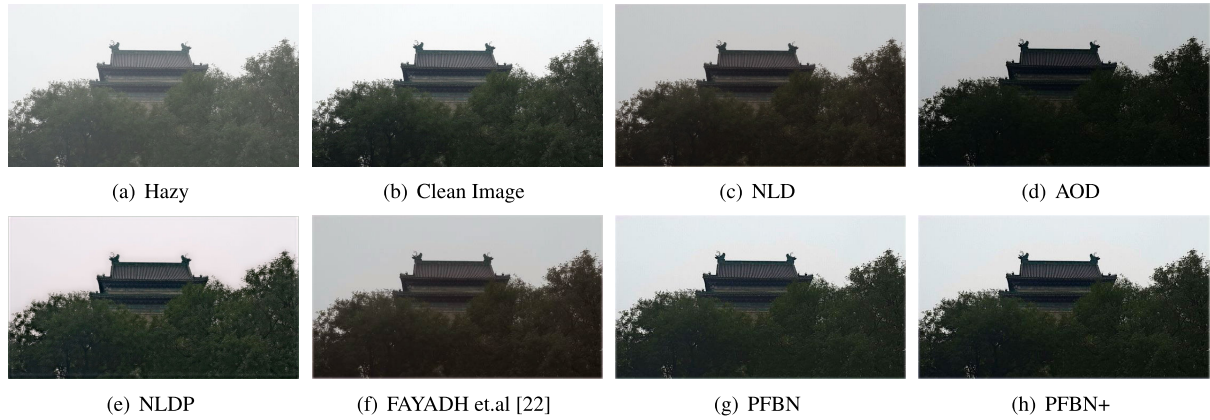


FIGURE 6. Visual quality comparison with state-of-the-arts. PFBN can reconstruct higher-quality detailed features than other state-of-the-arts, notes that the self-ensemble strategy can enhanced the performance slightly.

TABLE 1. Comparison of different loss functions.

	MSE loss	L1 loss	Negative SSIM	Recurrent MSE
PSNR	24.19	22.85	23.29	24.89
SSIM	0.972	0.939	0.958	0.980

3) STUDY OF LOSS FUNCTION

There are several widely adopted loss functions, such as MSE loss, L1 loss, negative SSIM loss and recurrent MSE (RMSE) loss, the different formulas of which are as follows:

$$\begin{aligned}
 MSE &= \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n [X(i, j) - Y(i, j)]^2 \\
 L1 &= \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n |[X(i, j) - Y(i, j)]| \\
 -SSIM &= -\frac{(2\mu_X\mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)} \\
 RMSE &= \frac{1}{T} \sum_{t=1}^T MSE_t \tag{8}
 \end{aligned}$$

where X, Y are the reconstruction image and ground truth, respectively; m, n denote the height and width, respectively; μ refers to the mean value of all pixel values of an image; σ refers to the deviation; and σ_{XY} denotes the covariance of images X and Y . More details about SSIM can be found in [21]. In recurrent MSE loss, each iteration is forced to reconstruct the dehazing image and calculate the MSE loss, and then, it must sum all iteration losses to obtain the recurrent loss.

We compare these formulas by setting $B = 3, T = 7$ and training different networks with respect to each loss. The results are shown in Table 1. It is clear that recurrent MSE loss is most suitable, so we choose it as our method’s loss function.

4) STUDY OF ACTIVATION FUNCTION

Similarly, now, we investigate the influence of the activation function; the alternative activation functions that are widely used in image reconstruction include RELU [23], leaky

TABLE 2. Comparison of various activation functions.

	RELU	Leaky RELU	GELU
PSNR	24.89	24.06	24.78
SSIM	0.980	0.971	0.979

ReLU [24], and Gaussian error linear units (GELU) [25]. The formula for the above activation functions are as follows:

$$\begin{aligned}
 RELU(x) &= \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \\
 LRELU(x) &= \begin{cases} x, & x > 0 \\ \frac{x}{a}, & x \leq 0 \end{cases} \\
 GELU(x) &= 0.5x(1 + \tanh[\sqrt{2/\pi}(x + 0.04715x^3)]) \tag{9}
 \end{aligned}$$

where a in leaky ReLU is a constant that satisfies $a \in [1, +\infty]$, and $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$.

A comparison of various activation functions is shown in Table 2. It is clearly shown that the network with RELU achieves the highest performance. Therefore, in our network, RELU is employed.

B. COMPARISON WITH THE STATE-OF-THE-ART METHODS

In this subsection, we evaluate our PFBN and other dehazing methods on three benchmark datasets, i.e., SOTS [12], I-HAZE [26], and O-HAZE [27]. It is notable that the networks are trained on 27,256 selected training pairs from NYU2 datasets [28]. The SOTS consists of 500 pairs of test images, I-HAZE consists of 35 image pairs that are hazy and haze-free images, and O-HAZE contains 45 different outdoor scenes.

Our PFBN is compared with several state-of-the-art dehazing methods, including NLD [7], AODNet [2], NLDN [29] and a Geometric-Pixel Guided CNN proposed by [22]. We reproduce these methods to compare. Besides, the proposed self-ensemble methods are implemented to enhance the reconstruction ability. The only difference between PFBN and PFBN+ is that PFBN+ is evaluated with the self-ensemble strategy. Table 3 illustrates the comparison results

TABLE 3. Comparison with state-of-the-art methods on three benchmark datasets.

	NLD [29]	AODNet [2]	NLDN [30]	FAYADH et.al [22]	PFBN	PFBN+
SOTS	19.25/0.898	18.29/0.856	20.42/0.923	22.14/0.951	24.89/0.980	25.22/0.988
I-HAZE	14.43/0.767	14.12/0.654	13.98/0.732	14.47/0.804	15.51/0.861	15.86/0.862
O-HAZE	16.78/0.653	15.98/0.585	15.03/0.539	16.62/0.893	18.47/0.893	18.71/0.901

TABLE 4. Comparison of trainable parameters, GPU memory and running time.

	NLD [29]	AODNet [2]	NLDN [30]	FAYADH et.al [22]	PFBN
Number of parameters	128548	107564	446438	102486	875682
GPU Memory (MiB)	1841	1277	4491	9471	8823
Running time (Seconds)	0.068	0.054	0.143	0.294	0.212

with state-of-the-art methods for three benchmark datasets. The evaluation metrics are PSNR/SSIM. The results significantly indicate that our PFBN has a stronger capability of reconstructing a clean image than other methods, and PFBN+ indeed benefits from such a simple strategy. According to [30], self-ensemble strategy actually decreases the variance in each pixel of hazy images and keeps its bias unchanged.

The visual results are shown in Fig. 6. It can be clearly observed from Fig.6(a) that haze exists over the image. The results of the comparison methods can remove this haze significantly but are not effective. Such visual results illustrate that the proposed PFBN can reconstruct high-quality clean images.

The most critical advantage of our model is its performance. The number of optimized parameters of PFBN and other state-of-the-art methods is shown in Table 4. Although the number of parameters, GPU memory and running time is a little bigger than other states-of-the-arts, PFBN have more powerful ability to reconstruct a haze-free scene.

Notes that self-ensemble strategy is only used in the test processing, it don't require any extra parameters, but it will increase the GPU memory in test processing and the running time is about 8 times that of PFBN.

V. CONCLUSION

In this paper, a progressive feedback network for single-image dehazing was proposed, which is easy to train and can achieve high performance for dehazing. The employed feedback connections can fuse high-level and low-level information through iterations, which enhances the representation ability of the network. In the training process, the sub-losses are tied to train the network according to the rules of the feedback mechanism. Additionally, a self-ensemble method tends to enhance the performance of the network. The extensive experimental results show that the proposed methods can reconstruct clearer dehazing images in some common benchmark datasets compared with state-of-the-art methods.

REFERENCES

- [1] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. ICCV*, 2017, pp. 2961–2969.
- [2] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "AOD-Net: All-in-one dehazing network," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 4770–4778.
- [3] R. Fattal, "Single image dehazing," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 1–9, 2008.
- [4] R. Hide, "Optics of the atmosphere: Scattering by molecules and particles," *Phys. Bull.*, vol. 28, no. 11, p. 521, Nov. 1977.
- [5] S. G. Narasimhan and S. K. Nayar, "Chromatic framework for vision in bad weather," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Jun. 2000, pp. 598–605.
- [6] S. G. Narasimhan and S. K. Nayar, "Vision and the atmosphere," *Int. J. Comput. Vis.*, vol. 48, no. 3, pp. 233–254, 2002.
- [7] D. Berman, T. Treibitz, and S. Avidan, "Non-local image dehazing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1674–1682.
- [8] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, "Single image dehazing via multi-scale convolutional neural networks," in *Proc. Eur. Conf. Comput. Vis. Amsterdam, The Netherlands: Springer*, Oct. 2016, pp. 154–169.
- [9] X. Jiang, J. Sun, C. Li, and H. Ding, "Video image defogging recognition based on recurrent neural network," *IEEE Trans. Ind. Informat.*, vol. 14, no. 7, pp. 3281–3288, Jul. 2018.
- [10] Y. Tan and G. Wang, "Image haze removal based on superpixels and Markov random field," *IEEE Access*, vol. 8, pp. 60728–60736, 2020.
- [11] A. R. Zamir, T.-L. Wu, L. Sun, W. B. Shen, B. E. Shi, J. Malik, and S. Savarese, "Feedback networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 1308–1317.
- [12] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang, "Benchmarking single image dehazing and beyond," 2017, *arXiv:1712.04143*.
- [13] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [14] F. Shama, R. Mechrez, A. Shoshan, and L. Zelnik-Manor, "Adversarial feedback loop," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 3205–3214.
- [15] H. Yang, D. Yan, L. Zhang, D. Li, Y. Sun, S. You, and S. J. Maybank, "Feedback graph convolutional network for skeleton-based action recognition," 2020, *arXiv:2003.07564*.
- [16] J. Ding, H. Guo, H. Zhou, J. Yu, X. He, and B. Jiang, "Distributed feedback network for single-image deraining," *Inf. Sci.*, vol. 572, pp. 611–626, Sep. 2021.
- [17] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 2472–2481.
- [18] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 136–144.
- [19] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in PyTorch," in *Proc. 31st Conf. Neural Inf. Process. Syst. (NIPS)*, Long Beach, CA, USA, 2017.
- [20] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
- [21] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[22] F. S. Alenezi and S. Ganesan, "Geometric-pixel guided single-pass convolution neural network with graph cut for image dehazing," *IEEE Access*, vol. 9, pp. 29380–29391, 2021.

[23] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. ICML*, 2010, pp. 1–8.

[24] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. ICML*, 2013, vol. 30, no. 1, p. 3.

[25] D. Hendrycks and K. Gimpel, "Gaussian error linear units (GELUs)," 2016, *arXiv:1606.08415*.

[26] C. O. Ancuti, C. Ancuti, R. Timofte, and C. De Vleeschouwer, "I-HAZE: A dehazing benchmark with real hazy and haze-free indoor images," 2018, *arXiv:1804.05091*.

[27] C. O. Ancuti, C. Ancuti, R. Timofte, and C. De Vleeschouwer, "O-HAZE: A dehazing benchmark with real hazy and haze-free outdoor images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 754–762.

[28] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from RGBD images," in *Proc. Eur. Conf. Comput. Vis.* Florence, Italy: Springer, 2012, pp. 746–760.

[29] S. Zhang, F. He, and W. Ren, "Nldn: Non-local dehazing network for dense haze removal," *Neurocomputing*, vol. 410, pp. 363–373, Jan. 2020.

[30] J. Ding, D. Bao, Q. Wang, X. He, H. Bai, and S. Li, "A novel multi-dictionary framework with global sensing matrix design for compressed sensing," *Signal Process.*, vol. 152, pp. 69–78, Nov. 2018.



YUANHONG REN received the B.S. degree in information and computing science from the Wuhan University of Science and Technology and the M.S. degree in computer application technology from the Wuhan University of Technology. She is currently pursuing the Ph.D. degree with Donghua University. Her research interests include image processing and intelligent theory and control.



QI MAO received the Ph.D. degree from Donghua University, in 2019. He is currently an Associate Professor with the Shanghai University of Engineering Science. His research interests include artificial intelligence and medical image analysis.



TISONG LIANG was born in 1995. He is currently pursuing the M.S. degree with the Shanghai University of Engineering Science. His research interests include image processing and machine learning.



ZHIWEI LI received the M.S. degree in electronic science and technology from Wuhan University, in 2006, and the Ph.D. degree from the Huazhong University of Science and Technology, in 2016. He is currently an Associate Professor with the Shanghai University of Engineering Science. His research interests include intelligent theory and control, image processing, and optoelectronics.



WUNENG ZHOU received the M.S. degree from Central China Normal University, in 1982, and the Ph.D. degree in control science and engineering from Zhejiang University, in 2005. He is currently a Professor with Donghua University. His research interests include system theory and control theory, control engineering, and robust control theory and application.

...