



# Image Dehazing Based on Deep Multiscale Fusion Network and Continuous Memory Mechanism

Qiang Li, Zhihua Xie<sup>(✉)</sup>, Sha Zong, and Guodong Liu

Key Laboratory of Optic-Electronic and Communication, Jiangxi Science and Technology Normal University, Nanchang, China  
xie\_zhihua68@aliyun.com

**Abstract.** This paper proposed a multi-scale feature fusion image dehazing network by incorporating a contiguous memory mechanism (MFFDN-CM). Specifically, the pixel attention mechanism, continuous memory strategy and residual dense blocks are integrated into the dehazing model with a prevalent encoder-decoder structure (U-Net). Firstly, our model obtains multiscale feature maps by subsampling operations, and further employs skip connections between the corresponding network layers to connect the feature maps between the encoder and the decoder for good feature fusion. Then, we introduce a continuous memory residual block to strengthen the information flows for feature reuse. Moreover, to leverage detail representation and accomplish adaptive dehazing according to the haze density, MFFDN-CM adopts a pixel attention module on the skip connections to combine the residual dense block module of the corresponding decoding layers. Finally, multiple residual blocks are exploited on the bottleneck in encoder-decoder structure to prevent network performance degradation due to vanishing gradients. Experimental results demonstrate the proposed model can achieve better hazing performance than the state-of-the-art methods based on deep neural network.

**Keywords:** Image dehazing · Attention mechanism · Continuous memory mechanism · Residual dense block module · U-Net

## 1 Introduction

Scene detection and enhancement under special weather have become a representative task for robust machine vision systems. As shown in Fig. 1, images captured in hazy days are easily affected by fog or haze, resulting in blurred image details, low contrast, and loss of meaningful textures. To solve these problems, image dehazing algorithms emerge as a solution to improve the quality of recorded images. Aiming to obtain high-quality images, it is necessary to perform dehazing restoration on the images. Following this effort, image dehazing has been explored for many years [2, 4–6] and many algorithms has been developed for image dehazing. To the best of our knowledge, these researches

usually simulate a hazy image by atmospheric scattering and attenuation model, as formulated in (1).

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

where  $x$  is the pixel position,  $J(x)$  is the latent haze-free image,  $t(x)$  is the medium transmission map, and  $A$  is the global atmospheric light which indicates the intensity of ambient light.



**Fig. 1.** Examples of realistic hazy images

Based on the physical model, given a captured hazy image, the haze-free image can be generated by estimating the transmission map and global atmospheric light values. To recover haze-free images  $J(x)$  from hazy images  $I(x)$ , end-to-end deep learning methods have shown their promising capability. Some early methods [1, 2] tried to use convolutional neural networks (CNN) to predict the transmission map, and then other methods [3] were used to estimate the value of atmospheric light. However, the estimation of transmission maps or atmospheric light values from a single hazy input is not an easy task, due to the uncertainty of atmospheric light values [7] and the difficulty of capturing ground truth data of transmission maps. Furthermore, inaccurate estimates of transmission maps or atmospheric light values can seriously affect the recovery of haze-free images. To solve this problem, several algorithms [8–10] estimate latent images directly or iteratively [11] based on deep CNNs. However, these methods mainly adopt universal network architectures (e.g., GDN [12], PFFNet [13], GFN [14], GCANet [8]), which are not well optimized for some typical problems that arise during image dehazing. Although these networks can represent multi-scale features, as the network goes deeper, spatial information will be lost, thereby reducing the feature representation capability of the network. It is also worth noting that most image dehazing networks treat the characteristics of pixels equally, which is unsuitable for dealing with images with uneven distribution of haze, and it is prone to appear the phenomenon of uneven image dehazing or artifacts. In response to the above problems, we creatively add residual dense module [15] and pixel attention [16] to the multi-scale network to enrich feature representation and guide the network to process different feature information adaptively. Following this motivation, this paper proposes a multi-scale feature fusion image dehazing deep model with a continuous memory mechanism (MFFDN-CM), which is achieved by aggregating residual dense modules and pixel attention modules into U-Net. In summary, the main contributions are as follows:

- (1) We introduce Residual Dense Block (RDB) module to extract hierarchical features from convolutional layers via local and global feature fusion for stable training and preservation of global features. Moreover, because of information flow interaction between RDB modules, all information is fully utilized through global residual learning.
- (2) Considering the uneven distribution of haze on different image pixels, we integrate a pixel attention module (PA) to make the network pay more attention to useful features, such as thick haze pixels and high-frequency image regions, thereby achieving adaptive dehazing operation.
- (3) Inspired by [17], we incorporate continuous memory residual blocks into the encoder and decoder, which not only can reduce the occupied memory and training time but also increase the information flows through feature reuse, aiming at further improving the network feature representation ability.

Experimental results show that our proposed method can achieve the best performance on the benchmark SOTS dataset [18].

## 2 Related Works

Recent years witnessed the dramatically progress on image dehazing researches. Generally speaking, dehazing methods can be roughly divided into two categories: traditional feature engineer-based algorithms and deep learning-based methods.

As for traditional feature engineer-based algorithms, image dehazing mostly explore hand-crafted features to predict haze-free images. These features are mainly extracted based on some image priors or assumptions [3]. It can be found that most patches in realistic images contain several pixels with very low brightness which is close to zero in at least a color channel, called the dark channel prior (DCP). Early methods used DCP as the original image before estimating the transmission map of image dehazing, and obtained relatively good results. However, there are obvious artifacts in the output dehazed image due to inaccurate estimation of the transmission map. Subsequently, many works have been devoted to improve DCP methods via using boundary constraints [19] or non-local total variation [20]. Although the artifact problem is resolved, these methods lead to severe color distortion in the sky area, since the DCP assumption is based on the outdoor clear image that excludes the sky area. Briefly, the traditional methods based on priors are not robust against diverse scenes.

On the other hand, with the advance in convolutional neural networks (CNNs) and the availability of large-scale synthetic datasets, deep learning-based approaches for image dehazing have received significant attention in recent years. MSCNN [1] and DehazeNet [2] are the earliest methods to utilize CNN for image dehazing. Both methods use a CNN to estimate the transmission map, which is then used to obtain a haze-free image based on Eq. (1). In addition, there is AOD [6], which uses traditional prior methods to reformulate Eq. (1) to build an end-to-end model. Whereas, Gated Fusion Network (GFN) [14] derives three augmented versions from hazy images, then treats them as inputs and performs gated fusion on their results. UR-Net [21] is a recently proposed model, which is based on an encoder-decoder network with bottleneck residual

blocks. The model achieves good results on some standard dehazing datasets. Shao et al. [27] developed a domain adaptation network (DA) with two image translation modules between synthesized and real hazy images and two image dehazing modules to alleviate the domain shift problem. However, the image dehazing modules of DA [27] learned CNN features from input hazy image to predict only one factor (i.e., the latent haze-free image), thereby hindering the dehazing performance.

Although the existing single-image dehazing algorithms can remove the haze from hazed images to a certain extent, it is difficult to obtain the perfect high-quality haze-free image. To be specific, compared with several existing state-of-the-art single-image dehazing methods, PFFNet [13] exhibits ideal dehazing performance on the RESIDE [18] dataset, which can recover most of the outdoor haze images as expected. Meanwhile, it will produce certain artifacts and color distortions when applied to indoor haze images. To address the poor performance of PFFNet on indoor hazy images, the main motivation of this paper is focused on improvements on the PFFNet baseline to make our method perform better on indoor and outdoor haze images.

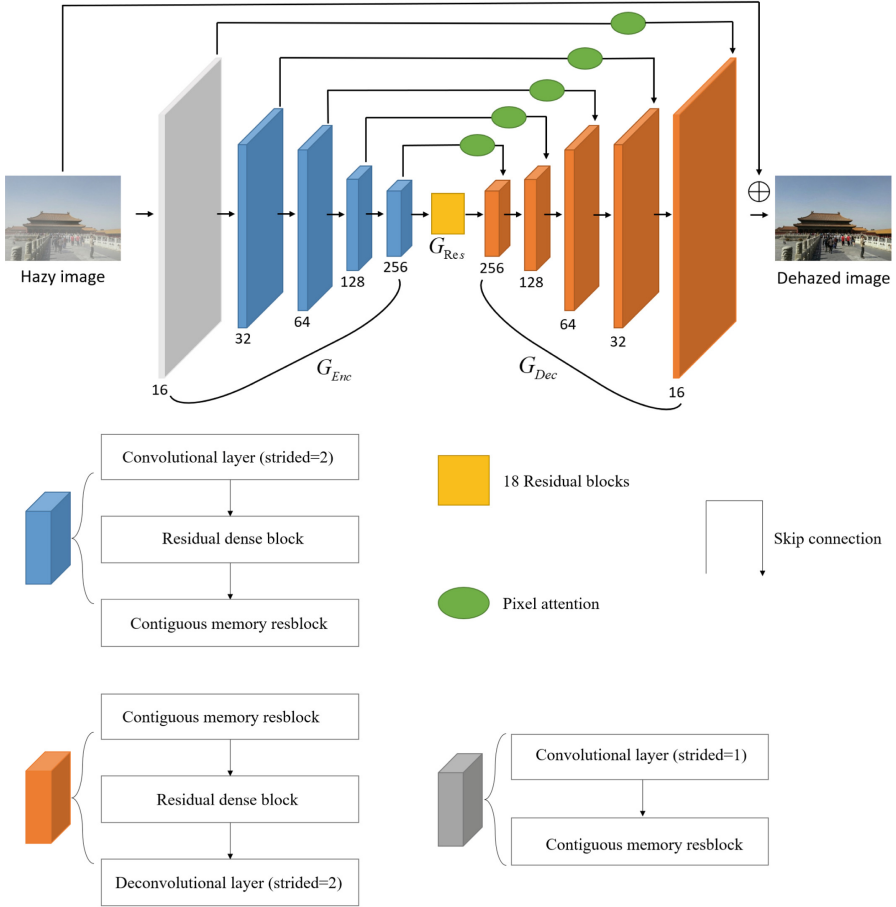
### 3 Proposed Dehazing Method

In this paper, we propose a multi-scale feature fusion image dehazing network incorporating a contiguous memory mechanism (MFFDN-CM). This is an end-to-end trainable CNN model, which develops a deep model equipped with multi-scale feature fusion driven by continuous memory, attentional module and residual-dense block for image dehazing.

#### 3.1 Network Architecture

Figure 2 shows the entire architecture of the proposed MFFDN-CM. As can be seen from Fig. 2, MFFDN-CM is evolved from U-Net [22] and ResNet [23]. Specifically, MFFDN-CM inherits the basic encoder-decoder structure of U-Net, and we embed 18 residual blocks (resblocks) at the junction of the encoder module and the decoder module (i.e., the bottleneck structure). In addition, MFFDN-CM inherits the main idea of ResNet (that is, learning the residual image instead of directly learning the image dehazing result), and adds the residual image and the original haze image to obtain the final image dehazing result. MFFDN-CM consists of four parts: an encoder module ( $G_{Enc}$ ), a bottleneck recovery module ( $G_{Res}$ ), Pixel attention based skip connection module (PA), and a decoder module ( $G_{Dec}$ ). Specifically, the goal of the encoder module is to extract shallow feature maps of images and enhance the preservation of local information through convolution and residual dense blocks. The decoder module obtains the depth feature map of the image through deconvolution. Considering that subsampling shallow features will be partially lost, this paper adds continuous memory residual blocks in both the encoder and the decoder, which can increase the information flow through feature reuse and further enhance the feature expression ability of the network.

Then, a skip connection module based on pixel attention (PA) is used to assign weights to the input feature maps according to their importance, so that the fusion of the shallow feature map and the deep feature map is more effective. In addition,



**Fig. 2.** The schematic illustration of the MFFDN-CM

the bottleneck recovery structure consists of 18 residual blocks to prevent the network performance degradation caused by vanishing gradients. Simply speaking, MFFDN-CM adopts an encoder-decoder structure to learn residual images, and adds the original haze image to the residual image to generate the final image dehazing result. Furthermore, the activation functions and normalization operations used in MFFDN-CM are Parametric Rectified Linear Unit (PReLU) [24] and Group Normalization (GN) [25].

### 3.2 Continuous Memory Residual Block

Inspired by [17], a continuous memory residual block is utilized in both the encoder and the decoder, which can increase the information flow through feature reuse and further enhance the feature expression ability of the network. The Continuous Memory Residual block (CMres) consists of two ordinary residual blocks (with a kernel size of  $3 \times 3$ ) and a convolutional layer (with a kernel size of  $1 \times 1$ ). As shown in Fig. 3, a contiguous

memory mechanism is realized by the operation similar to dense block which increasing the information flow through feature reuse. To reduce memory usage and runtime, connections are only used between each normal resblock rather than each convolutional layer. Let  $F_{n-1}$  and  $F_n$  be the input and output of the  $n$ th CMres respectively, then the output of the  $n$ th CMres can be expressed as

$$F_n = H_n[F_{n-1}, F_{n,1}, F_{n,2}] \tag{2}$$

where  $H$  is the bottleneck layer.  $F_{n,1}$  and  $F_{n,2}$  are the feature maps generated by ordinary residual blocks 1st and 2nd on  $n$ th layers.

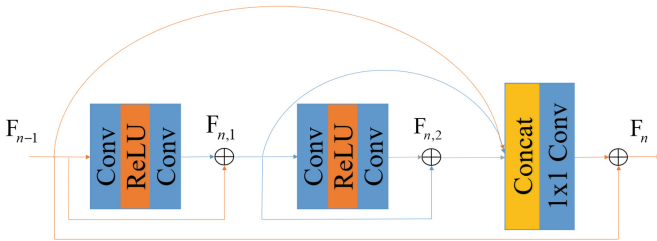


Fig. 3. Contiguous memory residual blocks

### 3.3 Pixel Attention Module

According to [16], we adopted a pixel attention (PA) module in skip connection, which assigned weight to each pixel of the shallow feature map that still retained the texture information of the original image, so that the final restored image was closer to the pixel distribution of the original image in detail texture. In addition, the distribution of haze on each pixel is usually uneven. The addition of pixel attention module also makes the network model pay more attention to the density information of haze, providing flexibility for the model to deal with haze of different density.

As shown in Fig. 4, we directly feed the input  $F$  (the output of the previous layer) into two convolutional layers with ReLU and sigmoid activation functions. The shape changes from  $C \times H \times W$  to  $1 \times H \times W$ .

$$PA = \sigma(Conv(\delta(Conv(F)))) \tag{3}$$

where  $\sigma$  is the sigmoid function,  $\delta$  is the ReLU function, and  $F^*$  is the output of the pixel attention module. Finally, we perform element-wise multiplication of the weights of the input  $F$  and pixel  $PA$ .

$$F^* = F \otimes PA \tag{4}$$

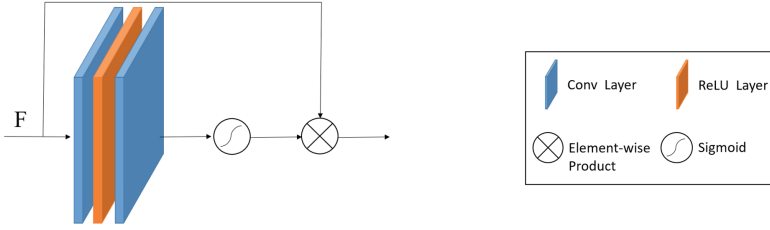


Fig. 4. Pixel attention module

### 3.4 Residual Dense Block (RDB)

To take full advantage of the hierarchical nature of all convolutional layers, as shown in Fig. 5, we employ [15] Residual Dense Blocks (RDB) to extract rich local features by densely connecting convolutional layers. RDB can realize the direct connection from the previous RDB state to all layers of the current RDB, and then use the local feature fusion of RDB to adaptively learn more effective feature information from the previous and current local features, which is conducive to stability wider network training and improved network stability and performance. The final output of the  $d$ -th RDB can be formulated as

$$F_d = F_{d-1} + F_{d,LF} \tag{5}$$

This module introduces a  $1 \times 1$  convolutional layer to adaptively control the output information. We name this operation as local feature fusion formulated as

$$F_{d,LF} = H^d([F_{d-1}, F_{d,1}, \dots, F_{d,c}, \dots, F_{d,C}]) \tag{6}$$

where  $H^d$  denotes the function of the  $1 \times 1$  Conv layer in the  $d$ -th RDB.

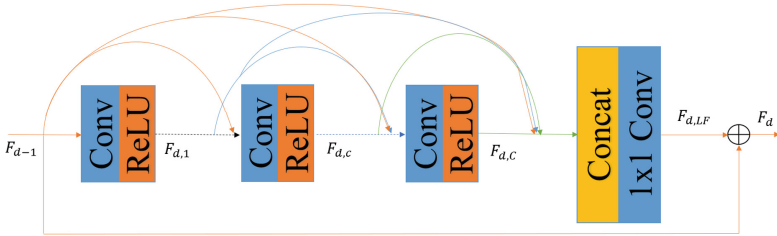


Fig. 5. Residual dense block

### 3.5 Loss Function

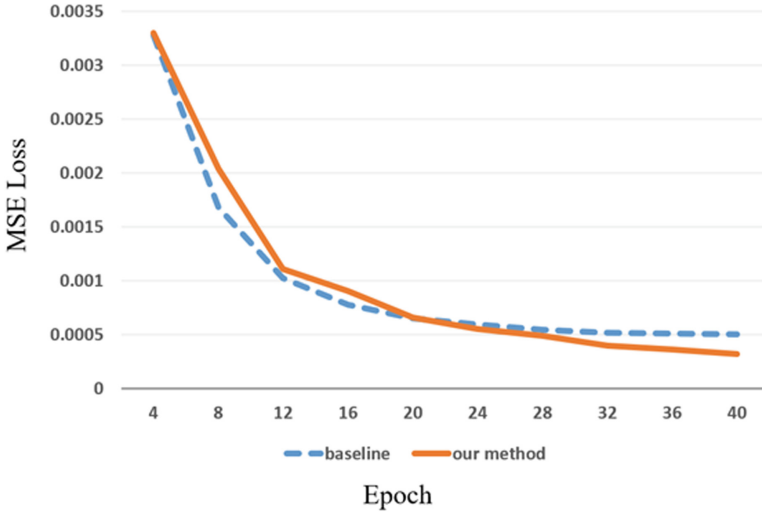
To effectively train the proposed network, MSE loss is exploited as the loss function, which is a pixel-based loss. MSE loss is defined as:

$$L_{MSE} = \frac{1}{N} \sum_{x=1}^N \sum_{i=1}^3 ||J_i(x) - j_i(x)||^2 \tag{7}$$

where  $J_i(x)$  and  $j_i(x)$  are the value of pixel  $x$  in the  $i$ -th color channel in the output dehazed image and the ground truth respectively, and  $N$  is the total number of pixels.

## 4 Experimental Results and Analysis

### 4.1 Implement Details



**Fig. 6.** Training loss curves of different methods

As shown in Fig. 2, the proposed network contains five hierarchical convolutional layers and five hierarchical deconvolutional layers. According to [13] 18 residual blocks are also used in our bottleneck recovery module ( $G_{Res}$ ). In the first convolutional layer of the encoder module, the filter size is set to  $11 \times 11$  pixels, and in all other convolutional and deconvolutional layers, the filter size is  $3 \times 3$ . We jointly train the MFFDN and CM modules and use mean squared error (MSE) as the loss function to constrain the network output and ground truth. The entire training process consists of 40 epochs optimized by the ADAM solver [26] with  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ , and a batch size of 10. The initial learning rate is set to 0.0001, and the decay rate is 0.5 after every 10 epochs. All experiments are performed on NVIDIA GeForce RTX 3070 GPU. The training loss curves are shown in Fig. 6. It can be found that our model has better performance in terms of loss convergence.



## 4.2 Dataset

To learn a general dehazing model for indoor and outdoor scenes, we select 9000 outdoor hazy/clean image pairs and 7000 indoor hazy/clean image pairs as training from the resident training dataset [18] by removing redundant images in the same scene set. To further expand the training data, we randomly flip the images horizontally or vertically. To evaluate the effectiveness, we use the synthetic objective test set (SOTS) dataset, which contains 500 indoor and 500 outdoor images. For comparison, all methods are trained on selected resident training datasets and evaluated on SOTS.

## 4.3 Evaluation Results

### Quantitative Results

We compare our method with single-image dehazing methods based on handcrafted features (DCP [3]) and deep convolutional neural networks (AOD [6], GFN [14], GCANet [8], GDN [12], DA [27], DMT [28] and PFFNet [13]). We use prevalent PSNR and SSIM metrics to evaluate the quality of each restored image. Higher PSNR and SSIM values indicate a better image dehazing effect. Since most existing deep model-based methods are trained on various datasets, for fair comparison, we retrained GFN [14], GDN [12], DA [27], DMT [28], GCANet [8] and PFFNet [13] model on the same training dataset. The quantitative results on the outdoor sets and the indoor sets of the SOTS database are listed in Table 1 and Table 2, respectively.

In these two tables, the metrics indicating the best performance are marked in bold. Table 1 lists the average PSNR and SSIM values of the image dehazing results obtained by applying each method to all the outdoor hazy images from the SOTS dataset. Table 2 lists the average PSNR and SSIM values obtained by applying each method to all the indoor hazy images from SOTS.

From Table 1, the proposed MFFDN-CM obtains the best outdoor image dehazing effect with the highest PSNR value of 34.30 and the highest SSIM value of 0.984. Both the PSNR and SSIM values obtained by the MFFDN-CM are higher than those obtained by other eight methods. For the eight reference methods, DMT obtains the second-highest PSNR and GDN gets the second-highest SSIM index. DCP and AOD obtain the lower PSNR and SSIM values indicating the worst image dehazing effect. Meanwhile, GFN, GCANet, DA and PFFNet obtain intermediate image dehazing results with middle PSNR and SSIM values. Moreover, it can be seen in Table 2 that the MFFDN-CM also obtains the highest PSNR and SSIM values (i.e., 32.98 and 0.980), indicating that it offers the best image dehazing effect on the indoor hazy images. Both DCP and AOD obtains the worst image dehazing effect with the lowest PSNR and SSIM. GDN obtains the second highest PSNR value and the second-highest SSIM value. GCANet obtains the third-highest PSNR value. DMT and PFFNet obtains an intermediate image dehazing effect. In a word, the experimental results on the SOTS dataset validate that MFFDN-CM obtains the best image dehazing effect among the state-of-the-art methods.

**Table 1.** Quantitative evaluations on the outdoor sets of the SOTS dataset

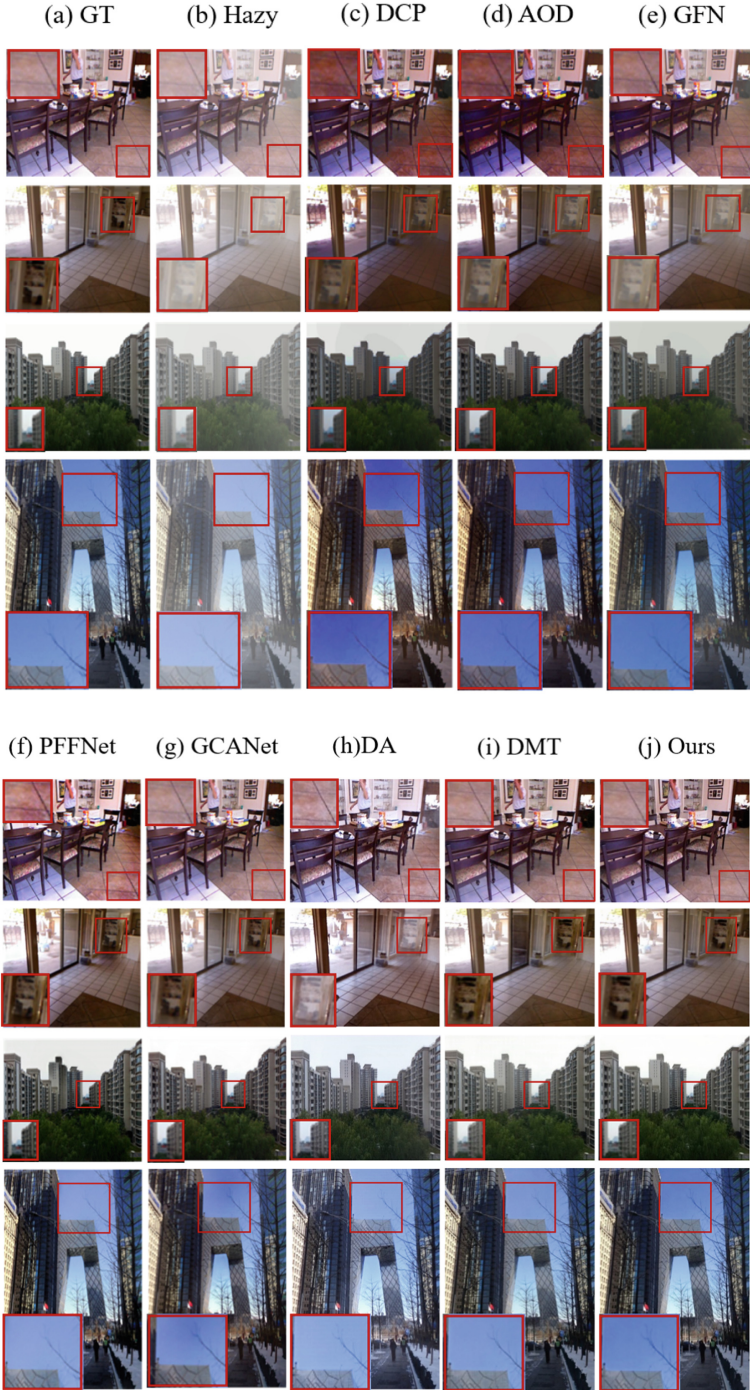
Methods	DCP	AOD	GFN	GCANet	PFFNet	GDN	DA	DMT	Ours
PSNR	17.54	19.74	25.02	26.20	31.11	30.86	28.48	31.22	<b>34.30</b>
SSIM	0.848	0.874	0.916	0.916	0.968	0.982	0.942	0.971	<b>0.984</b>

**Table 2.** Quantitative evaluations on the indoor sets of the SOTS dataset

Methods	DCP	AOD	GFN	GCANet	PFFNet	GDN	DA	DMT	Ours
PSNR	19.96	17.86	23.20	30.06	27.32	32.16	26.68	28.92	<b>32.98</b>
SSIM	0.870	0.794	0.882	0.916	0.940	0.978	0.922	0.951	<b>0.980</b>

### Qualitative Results

To further qualitatively compare the image dehazing effect of different methods, Fig. 7 exhibit visual image dehazing results on several representative synthetic indoor hazy images, synthetic outdoor hazy images. In detail, Fig. 7 show image dehazing results of applying eight methods (i.e., DCP [3], AOD [6], GFN [14], GCANet [8], PFFNet [13], DA [27], DMT [28], and MFFDN-CM) to four synthetic hazy images from the indoor subset and the outdoor subset of SOTS, respectively. In Fig. 7, GT is the ground truth image and Hazy is the hazing image. To be convenient to compare and analysis, the places with significant differences after dehazing by different methods are framed in red. It can be seen from Fig. 7 that the indoor image dehazing results obtained by GFN and AOD have obvious residual haze. Meanwhile, the indoor image dehazing results obtained by DCP has obvious color distortion. The indoor image dehazing results about the second line of Fig. 7 (f–j) obtained by DA retain fewer residual haze than GFN and AOD. More importantly, the image dehazing results of indoor obtained by GFN, PFFNet, DA and AOD have various degrees of color distortion. The proposed MFFDN-CM achieves the best indoor image dehazing performance, since our dehazing images are more similar to the ground truths (i.e., haze-free images). As a result, qualitatively demonstrated from the visual aspect, our proposed method significantly improves the indoor dehazing performance compared to the baseline PFFNet. On the other hand, we can draw the consent conclusion from the image dehazing results on the two outdoor hazy images shown in Fig. 7. The first five methods fail to completely remove image haze, especially DCP and AOD. Besides, GCANet removes haze, but generates an unnatural image background such as the sky in Fig. 7(g). Additionally, the approach based on DMT performs well to some extent, while the brightness distortion occurs in the sky area of the dehazed images. In short, MFFDN-CM obtains the image dehazing results closest to the ground truths, thus showing the super dehazing efficacy.



**Fig. 7.** Qualitative comparisons with state-of-the-art methods on SOTS

#### 4.4 Ablation Study

In order to figure out the contributions of different components to the dehazing performance, we conduct ablation studies by training several variants of the proposed network on the RESIDE dataset. These variants include baseline PFFNet, PFFNet with CMres module (Modle-1), PFFNet with PA module (Model-2), PFFNet with RDB module (Model-3), and our proposed network MFFDN-CM (Modle-4). The ablation results are reported in Table 3. According to the dehazing performance on the SOTS indoor dataset from Table 3, MFFDN-CM outperforms other methods, confirming the importance of the roles played by the CMres module, RDB module and PA module respectively. Moreover, the CMres component contributes to the better improvement in comparison with the other components.

**Table 3.** Ablation results on different components in the proposed network. All the models are trained on the RESIDE dataset with the same hyper-parameters

Modules	Baseline	Modle-1	Modle-2	Modle-3	Modle-4
CMres		√			√
PA			√		√
RDB				√	√
PSNR	27.32	31.56	28.12	30.36	32.98
SSIM	0.940	0.971	0.965	0.967	0.980

## 5 Conclusions

In this paper, we propose a U-Net like multi-scale deep neural network for single image dehazing, which aggregates continuous memory residual blocks and residual dense modules and pixel attention modules. MFFDN-CM uses a pixel attention module at the skip connection, by combining the residual dense block module of the corresponding decoding layers, it helps to extract detail information fully and accomplish adaptive dehazing according to the haze density. Our encoder-decoder network enhances the features for dehazing and refines the dehazing results using a new continuous memory mechanism in the encoder and the decoder. This design not only enables spatial consistency but also reduces information dilution by feature reuse. Extensive ablation studies have verified the effectiveness of the incorporated components in our dehazing network. Experiment results on synthetic images also indicate that our method outperforms state-of-the-art methods quantitatively and qualitatively. Since it is difficult to obtain hazy/clean image pairs in the real world, unsupervised/self-supervised learning methods may be more suitable for image dehazing, and we will focus on this topic in the future.

**Acknowledgements.** This paper is supported by the National Nature Science Foundation of China (No. 61861020), the Jiangxi Province Graduate Innovation Special Fund Project (No. YC2021-X06) and the Nanchang Educational Big Data & Intelligent Technology Key Laboratory (No. 2020NCZDSY-012).

## References

1. Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., Yang, M.-H.: Single image dehazing via multi-scale convolutional neural networks. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9906, pp. 154–169. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-46475-6\\_10](https://doi.org/10.1007/978-3-319-46475-6_10)
2. Cai, B., et al.: DehazeNet: an end-to-end system for single image haze removal. *IEEE Trans. Image Process.* **25**(11), 5187–5198 (2016)
3. He, K., Jian, S., Tang, X.: Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(12), 2341–2353 (2010)
4. Zhang, H., Patel, M.V.: Densely connected pyramid dehazing network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2018)
5. Yang, D., Sun, J.: Proximal Dehaze-Net: a prior learning-based deep network for single image dehazing. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11211, pp. 729–746. Springer, Cham (2018). [https://doi.org/10.1007/978-3-030-01234-2\\_43](https://doi.org/10.1007/978-3-030-01234-2_43)
6. Li, B., et al.: AoD-Net: all-in-one dehazing network. In: Proceedings of the IEEE International Conference on Computer Vision (2017)
7. Li, Z., et al.: Simultaneous video defogging and stereo reconstruction. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2015)
8. Chen, D., et al.: Gated context aggregation network for image dehazing and deraining. In: IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE (2019)
9. Engin, D., Genç, A., Ekenel, H.K.: Cycle-Dehaze: enhanced cycleGAN for single image dehazing. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (2018)
10. Liu, X., et al.: Dual residual networks leveraging the potential of paired operations for image restoration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2019)
11. Chen, Y., et al.: Dual path networks. In: Advances in Neural Information Processing Systems, vol. 30 (2017)
12. Liu, X., et al.: GridDehazeNet: attention-based multi-scale network for image dehazing. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (2019)
13. Mei, K., Jiang, A., Li, J., Wang, M.: Progressive feature fusion network for realistic image dehazing. In: Jawahar, C.V., Li, H., Mori, G., Schindler, K. (eds.) ACCV 2018. LNCS, vol. 11361, pp. 203–215. Springer, Cham (2019). [https://doi.org/10.1007/978-3-030-20887-5\\_13](https://doi.org/10.1007/978-3-030-20887-5_13)
14. Ren, W., et al.: Gated fusion network for single image dehazing. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2018)
15. Zhang, Y., et al.: Residual dense network for image super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2018)
16. Qin, X., et al.: FFA-Net: feature fusion attention network for single image dehazing. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 7 (2020)
17. Xiang, L., Dong, H., Wang, F., Guo, Y., Ma, K.: Gated contiguous memory U-Net for single image dehazing. In: Gedeon, T., Wong, K.W., Lee, M. (eds.) ICONIP 2019. LNCS, vol. 11954, pp. 117–127. Springer, Cham (2019). [https://doi.org/10.1007/978-3-030-36711-4\\_11](https://doi.org/10.1007/978-3-030-36711-4_11)
18. Li, B., et al.: Benchmarking single-image dehazing and beyond. *IEEE Trans. Image Process.* **28**(1), 492–505 (2018)
19. Meng, G., et al.: Efficient image dehazing with boundary constraint and contextual regularization. In: Proceedings of the IEEE International Conference on Computer Vision (2013)

20. Liu, Q., et al.: Single image dehazing with depth-aware non-local total variation regularization. *IEEE Trans. Image Process.* **27**(10), 5178–5191 (2018)
21. Feng, T., et al.: URNet: a U-Net based residual network for image dehazing. *Appl. Soft. Comput.* **102**, 106884 (2021)
22. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015*. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
23. He, K., et al.: Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016)
24. He, K., et al.: Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: *Proceedings of the IEEE International Conference on Computer Vision* (2015)
25. Wu, Y., He, K.: Group normalization. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) *ECCV 2018*. LNCS, vol. 11217, pp. 3–19. Springer, Cham (2018). [https://doi.org/10.1007/978-3-030-01261-8\\_1](https://doi.org/10.1007/978-3-030-01261-8_1)
26. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. In: *International Conference on Learning Representations* (2015)
27. Shao, Y., et al.: Domain adaptation for image dehazing. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2020)
28. Liu, Y., et al.: From synthetic to real: image dehazing collaborating with unlabeled real data. In: *Proceedings of the 29th ACM International Conference on Multimedia* (2021)