ORIGINAL ARTICLE

DRDDN: dense residual and dilated dehazing network

Shengdong Zhang[1] · Jiaoting Zhang[3] · Fazhi He[2] · Neng Hou[4]

Abstract

Recently, deep convolutional neural networks (CNNs) have made great achievements in image restoration. However, there exists a large space to improve the performance of CNN-based dehazing model. In this paper, we design a fully end-to-end dehazing network, which can be called as dense residual and dilated dehazing network (DRDDN), for single image dehazing. In detail, a dilated densely connected block is designed to fully exploit multi-scale features through an adaptive learning process. The receptive field of the network is enlarged by dilation convolution without losing spatial information. Furthermore, we use deep residual to propagate the low-level features to high-level layers. Therefore, our model can fully exploit both low-level and high-level features for dehazing. Experiments on benchmark datasets and real-world hazy images show that the proposed DRDDN achieves favorable performance against the state-of-the-art methods.

Keywords Dense residual learning · Dehazing · Dilated densely connected block

1 Introduction

Haze or fog is a common phenomenon in the real-world scene, but it is usually undesirable in photographs. Users often want to extract the hidden clean image via removing haze from a hazy image. Removing haze can increase the contrast of an image and recover visual content with fine details. Therefore, dehazing is an attractive research topic in multimedia and vision communities. The researchers have taken a lot of attention to image dehazing in recent years due to its wide application in multimedia systems, such as image classification [1,2] and surveillance system [3]. All the above methods assume the inputs are clear, so restoring clean images from the inputs will boost the performance of computer vision systems.

Light obtained by the camera from an environment object is attenuated along the line of sight. Furthermore, the incom-

ing light is fused with the light, which is reflected into the line of sight by atmospheric particles (e.g., fog, haze and water droplets). This image degradation process can be mathematically formulated as [4]:

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

where the observed hazy image was represented by I, J denotes the corresponding haze-free scene radiance to be recovered, A is the global atmospheric light, and t is the transmission map defined as:

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

where β is the scattering coefficient of the atmosphere, $d(x)$ is the distance between the scene point and the camera. Based on the model (1), we can restore the haze-free image by:

$$J(x) = \frac{1}{t(x)}I(x) - A\frac{1}{t(x)} + A. \quad (3)$$

As there only exists one single hazy image, I is available as inputting, recovering the corresponding clean image J is a highly hard problem. In order to address this ill-posed problem, early traditional dehazing methods are proposed by using additional information or multiple images [5–7] to compute the transmission map. For example, depth is restored from multiple images captured in very dif-

✉ Fazhi He
    fzhe@whu.edu.cn

1   Computer Science Engineering Department, Shaoxing University, Shaoxing 312000, China

2   School of Computer Science, Wuhan University, Wuhan 430072, China

3   College of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China

4   School of Computer Science, Yangtze University, Jingzhou 434023, China

ferent weather conditions [6]. Polarization-based methods introduce more constrains via taking images with different degrees of polarization [5,7]. Although these methods can remove haze from natural hazy images, the additional requirement of input restricts the application of these methods.

Recently, single image haze removal has achieved significant improvements. Most existing methods compute transmission map and atmospheric light firstly and then remove haze by using the inversed model (3). Prior-based methods are proposed to compute the transmission map. For example, He et al. [8] found a dark channel prior (DCP) based on the statistical results of outdoor clean images. The DCP shows that most local outdoor clean image patches contain some pixels whose intensity is very low in one color channel at least and then use it to estimate the transmission map. However, it may fail for the scene which contains objects that are similar to the atmospheric light. Meng et al. [9] introduce a prior-boundary constrain to improve the dehazing performance. However, this prior has the same problem as dark channel prior. Fattal [10] utilizes the prior that pixels in a local clear image patch tend to show a line in RGB color space to recover the transmission map. However, this method needs care to select only reasonable patches that hold the assumption are considered. Berman et al. [11] discover haze-line prior based on that a clear image can be represented well via hundreds of color clusters. Although this method can handle general hazy images, it cannot deal with heavy hazy images and light hazy images with white object. Zhu et al. [12] design a linear model based on the observation that depth is correlated with color attenuation and then, solve the parameters of the model with a supervised mode.

All the afore-mentioned methods are based on hand-crafted designed features, and we find that these methods do not work well for some real cases since the assumptions on hazy image do not always hold. To overcome this drawback, some methods are proposed in machine learning framework. Cai et al. propose a CNN-based dehazing method which uses CNN and BReLU layer to obtain effective features to compute the transmission map [13]. Ren et al. propose to employ large-scale and small-scale networks to estimate transmission map [14,15]. Other methods [16,17] focus on atmospheric light estimating. All above methods obtain dehazed result via computing the atmospheric light and transmission map separately, which will generate sub-optimization results.

From the above analysis, we can understand that how to obtain effective features for removing haze from a single image is critical. To effectively extract the features from input hazy image, we propose a novel fully end-to-end dense residual and dilated dehazing network, i.e., DRDDN, via convolutional neural networks as [18–22]. DRDDN takes the whole hazy image as the input and outputs the corre-

sponding clean image directly via dense residual network, which incorporates global and local contextual information into dehazing. Dilation convolution has been used in semantic segmentation. However, it induces gridding artifacts [23]. To eliminate the gridding artifacts, we design a novel dilation scheme. Our contributions can be summarized as follows:

- We introduce a unified framework (DRDDN) for high-quality single image dehazing. The proposed network obtains the corresponding clean image from a hazy image directly, which can avoid the estimation of intermediate variables. The network makes full use of both global and local features to capture the global structure and local details jointly.
- We design a dense residual network by adding a few identity shortcuts. The proposed dense residual network can propagate the features of shallow layers to deep layers, which are critical for recovering the details of haze-free image. The use of a dense residual can ease the training of the network.
- We propose a dilated densely connected block (DDCB) to generate representative features, which are used to reconstruct the final dehazed result. DDCB increases the receptive field by using dilation convolution and exploits new features via dense connection. In addition, we design a dilation rate scheme to eliminate gridding artifacts, which is induced by dilation convolution.

2 Related work

2.1 Separately optimize methods

Most existing prior dehazing methods [8–12,24–27] employ different priors to compute the transmission maps. All these priors are based on single hand-crafted designed priors, so they cannot generalize well to all real cases. To avoid using hand-crafted priors, some learning-based methods are proposed to learn haze relevant features automatically [13,14, 28]. Apart from estimating transmission map, some other works focus on estimating atmospheric light [16]. If the computation of atmospheric light or the transmission map is not good enough, the final dehazed image usually contains color distortions or haze. Zhang et al. designed a dehazing model [28], which estimates transmission map and image detail jointly.

2.2 Jointly optimization methods

If the estimation of transmission map or atmospheric light is not good enough, the quality of final dehazed image will be affected. A natural way to address this problem is to recover haze-free images from hazy images directly. End-
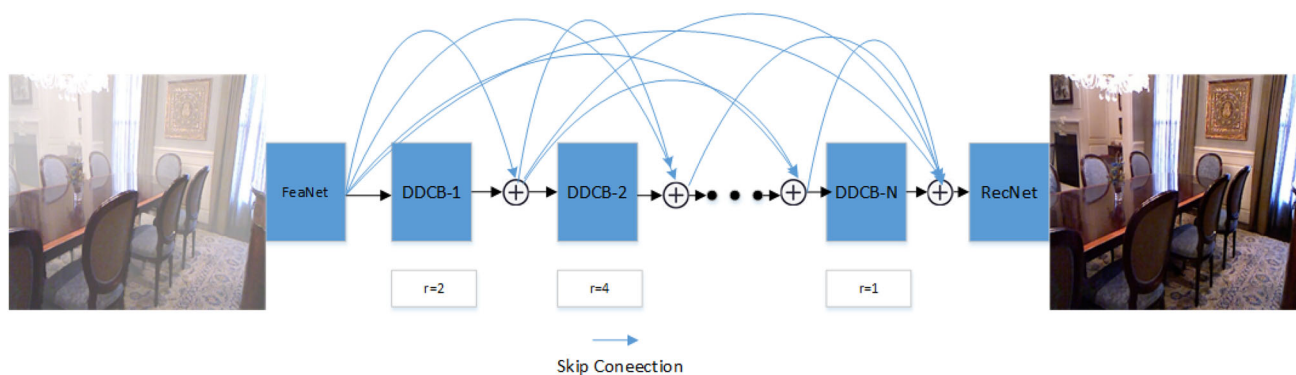
Fig. 1 Architecture of the proposed DRDDN. Residual learning is introduced to improve the information flow and make the network converge easily. Dilated densely connected block (DDCB) is introduced to enlarge the receptive field of the network and capture multi-scale features. Feature extraction layer extracts features from input hazy images, and the reconstruction layer recovers clean images from prior features extracted by DDCBs. Dilation rate is represented by r.

to-end CNN-based methods have been applied in image restoration [29–34] and saliency detection [35,36]; deep-learning-based methods [18,20,22,37–42] are also proposed to estimate clean image directly from the hazy input. Li et al. introduce an end-to-end method for dehazing [18] by fusing the transmission map and atmospheric light and formulating a new variable. Densely connected pyramid dehazing network (DCPDN) [21] is proposed by incorporating the atmospheric scattering model into the network. In [19], Ren et al. propose a network that firstly derives three inputs (the Contrast Enhancing, White Balance, and Gamma Correction) from an input hazy image and then, generates a haze-free image from these three inputs. FEED-Net [20] introduces a fully end-to-end dehazing method by employing dilation convolution. Liu et al. propose a GridDehazeNet [43], which combines multi-scale networks and attention for dehazing. Qin et al. proposed a dehazing network based on feature fusion attention [44]. To improve the dehazing performance of real haze images, Chen et al. proposed a dehazing method [45] based on well-trained backbone and come well-known physical priors. Zheng et al. proposed a fast dehazing method [46] for high-resolution hazy images. Although these methods are end-to-end fashion, they cannot deal with heavy hazy images. Zhang et al. designed a multi-scale dehazing network [47], which employs haze density map to remove haze more completely.

2.3 Convolutional neural networks (CNNs) for image tasks

Convolutional neural networks (CNNs) have been applied to image classification. Resnet [48] employs residual learning to train a very deep network, which improves the performance of classification. DenseNet [49] employs dense connection to concatenate the input features with the output features, enabling each layer to receive raw information from all previous layers. DenseNet shows that dense connections can be used to exploit new features. Fish et al. [23] employ dilation convolution to improve the performance of semantic segmentation. However, dilation convolution will cause gridding artifact.

3 Proposed method

This section introduces the details of our dense residual and dilated dehazing network which receives an original hazy image as an input and outputs the corresponding dehazed result. The methodology is to extract multi-scale features for different object sizes and employ dense residual learning to boost the image details. We refer to this network as DRDDN. As shown in Fig. 1, Our DRDDN mainly consists of four parts: Feature extraction layer, dilated densely connected blocks (DDCBs), reconstruction layer, and dense residual. Feature extraction layer is employed to extract shallow features from input hazy image. DDCBs are designed to extract hierarchical features from shallow features. Reconstruction layer is designed to reconstruct the clean image from hierarchical features. In what follows, we will introduce motivation, dense residual learning, and DDCB in detail.

3.1 Motivation

As stated in Sect. 1, CNNs have also been successfully applied in dehazing. However, we notice that all the models use limited contextual information due to the small receptive field size. In order to leverage the contextual information well, we employ the dilated convolution to enlarge the receptive field size of the proposed model. As shown in Fig. 2, objects tend to show different scales, which motivates us
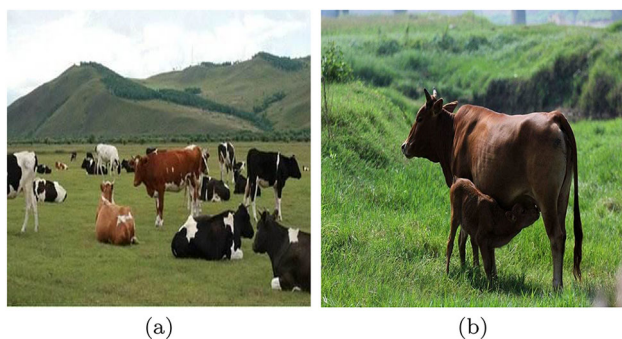
Fig. 2 Objects tend to show different scales. It is important to capture the objects using different receptive sizes
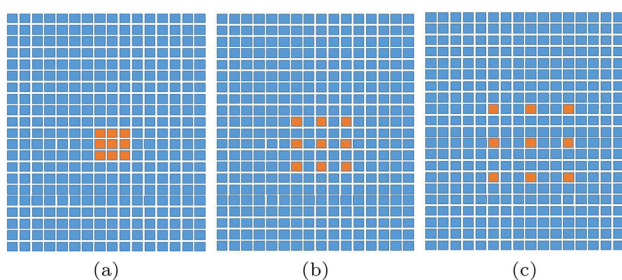
Fig. 3 Comparison of traditional convolution and dilation convolution. **a** The filter size is 3×3 with a dilation rate equals to 1 and covers 3×3 area in the features. **b** The filter size is 3×3 with a dilation rate equals to 2 and covers 5×5 area in the features. **c** The filter size is 3×3 with a dilation rate equals to 3 and covers 7×7 area in the features

to use dense connection to capture multi-scale objects. To capture the large object, we employ dilated convolution to increase the receptive of the model. In this paper, we propose a deep residual dilated densely connected dehazing network, which combines the advantages of DenseNet, dilation convolution and deep CNN. We propose a dilated densely connected block, which can be used to extract multi-scale features for dehazing. In addition, we add residual learning in the proposed model to further improve the dehazing performance. We also analyze why the dilation convolution can improve the receptive field of a model. As shown in Fig. 3, we can see that traditional convolution only covers a small area of input, whereas the dilation convolution can cover a larger area of input than traditional convolution. So the model with dilation can consider more contextual information. The proposed model can be used in denoising, deraining, super resolution, and style transfer.

3.2 Dense residual network

Our dense residual network (DRN) is designed to max the optimization ability of residual networks, which can ease the network training. Our dense residual network (DRN) contains long residual learning, short residual learning and other residual learning, which does not belong to long or short

residual learning. We now show more details about our proposed DRN structure (see Fig. 1), which contains N dilated densely connected blocks (DDCBs) and dense skip connections (DSC).

It has been demonstrated that global residual learning and stacked residual blocks can be used to construct deep CNN in deep residual network. Deep residual network with 1000 layers has shown its performance in visual recognition. However, very deep network constructed like this way would result in training difficulty and can hardly obtain more performance improvement than shallow networks for dehazing. This is the main reason why there lacks deep networks for single image dehazing. Inspired by ResNet and DenseNet, we propose a dense residual network. It has been proved that some of the middle and high-frequency information get lost at latter layers during a typical feedforward CNN process in very deep network, dense residual from previous DDCBs can avoid such loss and further boost high-frequency signals. The input of the first DDCB is the output of the feature extraction layer, which can be defined as F_0. Supposing that we have N DDCBs, we can define the output of Nth DDCB can be defined as F_n. The input of the first dilated dense connected block can be formulated as:

$$I_1 = F_0. \tag{4}$$

In order to pass on the information in shallow layers to deep ones, we design a dense residual strategy: the DDCB receives features from prior DDCBs. The input of the Nth DDCB can be defined as:

$$I_n = F_0 + \cdots + F_{n-1}, \tag{5}$$

where F_{n-1} represents the output of the $(N-1)$th dilated dense block. There are two reasons for the dense residual network can maximize the optimization ability of deep learning model. First, residual mapping has been proved to make the training of deep network easily, our dense residual network learns residuals between DDCBs. Second, ResNets uses local short connections to transmit information only between neighboring residual blocks. Long residual learning has been used to propagate the information from shallow layers to deep ones. DRN contains several direct connections and passes on information from the DDCB to subsequent ones, so all blocks can obtain information from the preceding DDCBs. As stated in DPN, the deep residual networks implicitly reuse the features through the residual path, which can pass the shallow layers features to deep layers and compensate mid/high-frequency information loss.

Gridding artifacts induced by dilation convolution have been discovered in semantic segmentation. Fish et al. [23] propose a method to remove the gridding artifacts. In this paper, we also develop a scheme for removing gridding arti-
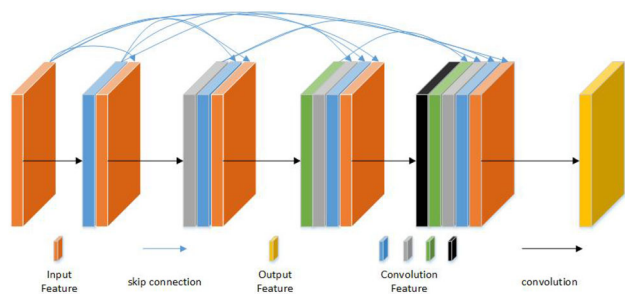
Fig. 4 Architecture of the proposed dilated densely connected block. DDCB contains dilation convolution layers and densely connections strategy, which is used to capture multi-scale features and maximize the information flow. The last convolution layer is used to reduce the channel number of features propagated to next DDCB. The dilation convolution layers in the same block share the same dilation rate

facts and enlarge the receptive field of the proposed model. The dilation rate of each block increases block by block until the dilation rate is 16, then decreases the dilation rate block by block until the dilation rate is 1. This scheme makes our model capture global structure and local information well and removes the gridding artifacts.

3.3 Dilated densely connected block

We introduce the details of the dilated densely connected block. Our DDCB contains dense connections and dilation convolution, which can be found in Fig. 4.

Densely connected strategy To further amplify the information flow between layers, we adopt the same strategy as DenseNet: the layer has direct connections to all subsequent layers. It has been proved that the dense connections block can improve the information flow and achieve a better convergence via connecting all layers. Consequently, we can express $F_{d,l}$ layer as

$$F_{d,l} = H_{dc}([F_{d,0}, \ldots, F_{d,l-1}]), \tag{6}$$

where H_{dc} represents the dilated convolution, $F_{d,c}$ denotes the output of cth dilated operation of dth dilated dense connect block.

Functions of the last dilation convolution layer in DDCB It is important to pass on the most representative features to subsequent layers. For this goal, we employ a dilation convolution layer to adaptively fuse the states from all preceding layers. As analyzed above, the feature maps of the prior layers are introduced directly to the last layer, it is helpful to reduce the feature number. Inspired by the MemNet, we introduce a 3×3 dilated convolution layer to output the most important information. Different from MemNet, our local features fusion can be incorporated into the DDCB with the same operation. The last dilation convolution layer is used to make the net-

work focus on more representative features and reduce the channels of features.

Larger receptive field It has been proved that the context information facilitates the reconstruction of the hazy pixels in image dehazing. Generally, prior works often enlarge the receptive field of CNN by increasing the filter size or the depth and downsampling. However, increasing the depth or filter size will not only increase the model size but also increase the computational time and resource [50]. Thus, existing CNN-based methods often use 3×3 filter with a large depth to increase receptive field of network. Furthermore, downsampling or progressive downsampling will cause loss of spatial information. The dehazing task becomes difficult when an object is not spatially dominant, for example, when the object is thin (e.g., leaves in the sky). In this case, the transmission or color cannot be recovered correctly. Our key idea is to keep spatial resolution in convolutional networks for dehazing. In order to keep the resolution and make a trade-off between the size of the receptive field and network depth, we use the recently proposed dilated convolution. The receptive field of our network with eight DDCBs is 365×365. If the traditional 3×3 convolution filters are used, the receptive field size of the proposed model is 69×69 with the same network depth (e.g., 34). Compared with traditional models, our model can capture large context information, which helps our model understand the scene configuration and recover the final clean image well.

Maximizing the information flow It has been proved that as the sampling rate becomes larger, the number of valid filter weights (i.e., the weights that are applied to the valid feature region, instead of padded zeros) becomes smaller [51]. In order to reduce the information loss in the calculation process, we use a dense connection to propagate the information from prior layers to subsequent ones. Residual learning can be used to propagate the information from prior blocks to subsequent ones. Combining densely connected and residual learning can maximize the information flow.

3.4 Loss function

Let $\mathcal{F}$ denotes the model which can be learned by the network, and Θ denotes the parameters of the proposed model. Let $\{I_i, i = 1, 2, \ldots, N\}$ and $\{J_i, i = 1, 2, \ldots, N\}$ denote the hazy input images and the corresponding clean ones, respectively. It has been widely proved that L_2 loss tends to obtain blurry dehazed results [21]. To solve this issue efficiently, we introduce a novel edge-preserving loss, which is composed of two different parts: L_1 loss and gradient loss, which can be defined as follows:

$$\mathcal{L}_e = \mathcal{L}_1 + \lambda_1 \mathcal{L}_g, \tag{7}$$

where $\mathcal{L}_e$ denotes the overall edge-preserving loss, $\mathcal{L}_1$ measures the content loss, $\mathcal{L}_g$ is the gradient loss, and λ_1 is used to control the importance of gradient loss. In addition, $\mathcal{L}_1$ is defined as follows:

$$\mathcal{L}_1 = \frac{1}{N} \sum_{i=1}^{N} \left\| \mathcal{F}(I_i, \Theta) - J_i \right\|_1, \tag{8}$$

where N is the number of training pair data and $\mathcal{L}_g$ is defined as:

$$\mathcal{L}_g = \frac{1}{N} \sum_{i=1}^{N} \left\| T(\mathcal{F}(I_i, \Theta), J_i) \right\|_1, \tag{9}$$

where $T(\mathcal{F}(I_i, \Theta), J_i)$ is the gradient difference between the predicted result and the ground truth, and it can be defined as:

$$
\begin{aligned}
T(\tilde{J}, \hat{J}) = \frac{1}{M} \sum_{i,j} & \left\| (\tilde{J}_{i,j-1} - \tilde{J}_{i,j}) - (\hat{J}_{i,j-1} - \hat{J}_{i,j}) \right\|_1 \\
& + \left\| (\tilde{J}_{i-1,j} - \tilde{J}_{i,j}) - (\hat{J}_{i-1,j} - \hat{J}_{i,j}) \right\|_1,
\end{aligned}
\tag{10}
$$

where $\tilde{J}$ represents the predicted dehazed result, $\hat{J}$ denotes the ground truth, M represents the number of pixels in the clean image J, and (i, j) denotes the location in the clean image J.

SSIM has been applied to evaluate the dehazing quality, and we also employ SSIM loss to further improve the dehazing quality. We define the SSIM loss as follows:

$$\mathcal{L}_{\text{ssim}} = \frac{1}{N} \sum_{i=1}^{N} ||1 - \text{ssim}(\mathcal{F}(I_i, \Theta), J_i)||_1. \tag{11}$$

Finally, we combine the edge-preserving loss and the SSIM loss to train the proposed network, which is defined as:

$$\mathcal{L}_o = \mathcal{L}_e + \lambda_2 \mathcal{L}_{\text{ssim}}, \tag{12}$$

where λ_2 is used to control the importance of the proposed SSIM loss.

4 Experimental results

In this section, we show the performance of the introduced method by comparing against the state-of-the-art methods including BCCR [9], DCP [8], NLD [11], MSCNN [14], DCPDN [21], GFN [19], DehazeNet [13], AOD-Net [18], and Proximal Dehaze-Net [52] on synthetic dataset and real-world hazy images quantitatively and qualitatively. We also

conduct the ablation study to evaluate the effectiveness of the proposed modules.

4.1 Synthetic dataset

As deep learning model requires a large training dataset to solve the dehazing problem, we synthesize a new hazy images dataset including both outdoor and indoor images to train the network. Similar to Ren et al.'s method [14], we also use NYU dataset [53] to synthesize indoor hazy images. In addition, the outdoor clean images are employed to synthesize outdoor hazy images. However, the outdoor images have no ground truth depth. To overcome this problem, we employ Liu et al.'s method [54] to obtain the depth map for each clean outdoor image. For a clean image J and corresponding depth, we synthesize a hazy image I according to the atmospheric scatter model (1). We can obtain the random atmospheric light $A = [n_1, n_2, n_3]$, where $n \in [0.8, 1.0]$ and we also generate a scatter coefficient $\beta \in [0.5, 2.0]$. Our dataset contains 2500 haze-free images, which includes 1400 indoor images and 1100 outdoor images. For each haze-free image, we generate 10 hazy images. The haze-free image and the corresponding hazy image are resized to the canonical size of 512×512 pixels.

4.2 Experimental settings

In the proposed DRDDN, we use 3×3 as the kernel size for all convolution layers. All the output channel number of the convolution layers (including dilated layers) is 24 except the last convolution layer, whose channel is 3. In our experiment, we set the number of DDCBs to 8. For each DDCB, we set the dilation rate to same for all dilated convolution layers. We set the dilation for all DDCBs to 2, 4, 8, 16, 8, 4 , 2, 1. All dilated layers are initialized using an identity initializer [55]. We set $\lambda_1 = 0.01$ and $\lambda_2 = 1$ in all the experiments. We employ the leaky rectified linear unit (LReLU) as the activation function of proposed model. We employ Adam optimizer with $\beta_1 = 0.9$ and $\beta_2 = 0.9999$ to train the network. The batch size and the learning rate are 1 and 0.0001, respectively. During training, we decrease the learning rate that decreases half for every 30 epochs. The network was trained for a total of 100 epochs by Tensorflow with a Nvidia GTX 1080Ti GPU. We incorporate the multi-scale in the training process by using different image sizes (300×300 and 600×600), which are resized from the training images.

4.3 Analysis

To show the improvement obtained by the modules introduced in the proposed network, we perform an ablation study that contains the following six experiments: (1) dilated densely connected structure (DDC) is a model without resid-

Table 1 Configurations of the proposed network

Layer	conv	DDCB1	DDCB2	DDCB3	DDCB4	DDCB5	DDCB6	DDCB7	DDCB8	conv
Channel	16	16	16	16	16	16	16	16	16	16
Output	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}
Dilation rate	1	2	4	8	16	8	4	2	1	1

Table 2 Configurations of the DDCB2

Layer	conv1	conv2	conv3	conv4	conv5
Size	3	3	3	3	3
Channel concat	16	16	16	16	16
		conv1	conv2, conv1	conv3, conv2, conv1	conv4, conv3, conv2, conv1
Dilation rate	4	4	4	4	4

Table 3 Configurations of the DDCB3

Layer	conv1	conv2	conv3	conv4	conv5
Size	3	3	3	3	3
Channel	16	16	16	16	16
concat		conv1	conv2, conv1	conv3, conv2, conv1	conv4, conv3, conv2, conv1
Dilation rate	8	8	8	8	8

Table 4 Configurations of the dense connection block (DCB)

Layer	conv1	conv2	conv3	conv4	conv5
Size	3	3	3	3	3
Channel	16	16	16	16	16
concat		conv1	conv2, conv1	conv3, conv2, conv1	conv4, conv3, conv2, conv1
Dilation rate	1	1	1	1	1

ual learning, (2) Deep residual densely connected structure (DRDC), (3) Deep residual dilated structure (DRD), (4) Deep residual dilated densely connected structure (DeRDDN), (5) Dense residual with 1×1 convolution for last layer in DDCB (1×1 conv), (6) Dense residual dilated densely connected structure (DRDDN) is our full model. The proposed model is presented more details in Tables 1, 2 and 3. The proposed model adds F_1 and F_2 as the input of DDCB2, which makes the model learning a dense residual for dehazing. DDCB (1×1 conv) changes the last layer in DDCB, and the new configuration is shown in Table 5. DeRDDN is a common residual learning and DDCB. DRD contains block without dense connections, we denote this block as DB, which is shown in Table 12. DRDC contains common residual learning and densely connected blocks (DCB), we show the DCB in Table 4. DDC contains no residual learning and DDCBs. All the designed models are trained with the same hyper-parameters as in Sect. 4.2 except the training epochs. All models are trained with 50 epochs. The eval-

uation is performed on the synthesized indoor hazy image dataset. In order to show the improvements, we synthesize a small hazy indoor dataset by using the pipeline proposed by Ren et al. [14]. This dataset contains 100 simulated hazy images and the corresponding haze-free images. Two criteria, such as PSNR and SSIM, are employed to show the dehazing performance. The SSIM and PSNR results averaged on dehazed images for the various configurations are tabulated in Table 6. As shown in Table 6, we can get the following conclusion: (1) The proposed method is able to keep the global structure for objects with relatively larger scale. (2) Dilation convolution can be used to improve the dehazing quality significantly. (3) Residual learning can help the converge of the model and improve the dehazing quality. (4) Dense connection can be used to maximize the information flow and improve the dehazing quality. (5) Dense residual can be used to propagate the information from shallow layers to deep layers and improve the dehazing quality. (6) We also note that fusing features by using 1×1 convolution

Table 5 Configurations of the DDCB (1 * 1)

Layer	conv1	conv2	conv3	conv4	conv5
Size	3	3	3	3	1
Channel	16	16	16	16	16
concat		conv1	conv2, conv1	conv3, conv2, conv1	conv4, conv3, conv2, conv1
Dilation rate	8	8	8	8	8

Table 6 Quantitative PSNR and SSIM results on the synthetic dataset

Metric	DDC	DRDC	DRD	DeRDNN	1×1 conv	DRDDN
PSNR	22.98	21.77	23.91	25.36	25.34	25.86
SSIM	0.90	0.89	0.92	0.93	0.93	0.94

Table 7 Average PSNR/SSIM of dehazed results on the SOTS dataset from **RESIDE** [56]

	DCP	BCCR	CAP	NLD	MSCNN	DehazeNet	AOD-Net	DCPDN	GFN	DDN	DRDDN
PSNR	16.62	16.88	19.05	17.29	17.57	21.14	19.06	15.86	22.30	25.17	**26.15**
SSIM	0.82	0.79	0.84	0.75	0.81	0.85	0.85	0.82	0.88	0.82	**0.97**

Bold indicates the best result

will reduce the dehazing performance. As shown in Table 6, the quantitative performance evaluated on synthesized indoor hazy image demonstrates the effectiveness of each module.

4.4 Quantitative evaluation

RESIDE dataset Synthetic objective testing set (SOTS) is a subset of REalistic Single Image DEhazing (RESIDE) [56]. Since the SOTS is simulated, the ground truth haze-free images are available and enabling us to test the dehazing performance qualitatively as well as quantitatively. We evaluate our algorithm on the SOTS and compare it with several state-of-the-art single image dehazing methods using PSNR and SSIM. As shown in Table 7, DDN [57] achieves a better dehazing performance than other state-of-the-art dehazing methods. The proposed method further boosts the dehazing performance.

Hazerd dataset Hazerd provides natural outdoor images with high accurate depth map, which enables us to better synthesize real outdoor hazy images taken under different hazy conditions. Proximal Dehaze-Net [52] provides a hazy dataset based on the hazerd. This dataset includes 128 images with different haze levels. Since all learning-based methods, such as DCPDN and DehazeNet, do not include images in Hazerd as training data, it is a fair way to compare them on Hazerd. In the experiments, we refer Proximal Dehaze-Net as PDNet. We compared the proposed model with state-of-the-art methods [9,11–14,18,19,21,58–61] and show the result in Table 8. As shown in Table 8, our deep learning-based method achieves the highest PNSR and SSIM on Hazerd and

exceeds the second best learning-based method PDNet [52] by 1.86 dB in terms of PSNR and 0.368 in terms of SSIM.

Evaluations on the O-HAZE dataset [63] Although we have evaluated the performance of the proposed method on RESIDE and Hazerd dataset, we note that the synthetic hazy images cannot consider the factors in haze formulation. The haze contained in synthetic hazy images may differ from the real haze. To overcome this issue, O-HAZE dataset is proposed by creating an air condition to generate haze. O-HAZE is a real hazy dataset. The haze in O-HAZE dataset is generated by two professional fog/haze machines. Compared with Hazerd and RESIDE, O-HAZE is more similar to real hazy image. O-HAZE provides 45 hazy images with corresponding haze-free images for various outdoor scenes. We compare our method with the state-of-the-art methods and report the performance by using PSNR and SSIM. From Table 9, we can see that our model achieves the best performance.

Evaluations on the Dense-HAZE dataset [64] Despite the dramatic increment in the interest shown for dehazing over the past decades of years, the evaluation of the dehazing methods remains not well studied, due to the absence of pairs of real hazy and corresponding haze-free reference images. To solve this issue, Dense-Haze was introduced. Dense-Haze contains 33 pairs of real hazy and the corresponding clean images of various scenes, which are dense and homogeneous. To show the high performance of the proposed method, we compare our method with some recently state-of-the-art methods [9,11,13,14,18,21,58,66–69]. The quantitative results in terms of PSNR and SSIM on the test set of the Dense-Haze dataset are reported in Table 10. As can be seen from Table 10, the proposed method ranks in the first place

Table 8 Average PSNR/SSIM of dehazed results on the Hazerd dataset [62]

	DCP	FVR	BCCR	CAP	NLD	GRM	MSCNN	DehazeNet	AOD-Net	DCPDN	GFN	PDNet	DRDDN
PSNR	17.66	16.17	16.31	18.56	18.82	17.47	19.10	19.53	18.13	18.82	19.18	20.14	**22.00**
SSIM	0.8430	0.8472	0.8337	0.8256	0.8355	0.7921	0.8540	0.8498	0.8266	0.8906	0.8601	0.8932	**0.9300**

Bold indicates the best result

Table 9 Average PSNR/SSIM of dehazed results on the O-HAZE dataset [63]

	DCP	FVR	BCCR	CAP	NLD	GRM	MSCNN	DehazeNet	AOD-Net	DCPDN	GFN	PDNet	DRDDN
PSNR	16.59	14.92	17.44	17.62	16.61	15.62	19.07	16.21	16.72	15.62	18.30	18.52	**20.10**
SSIM	0.74	0.70	0.75	0.71	0.75	0.58	0.77	0.67	0.68	0.62	0.72	0.75	**0.78**
CIEDE2000	20.75	19.59	22.73	11.22	15.67	17.97	14.67	17.34	19.11	22.31	17.23	17.40	**10.53**

Bold indicates the best result

SSIM and PSNR—higher is better and CIEDE2000—lower is better

Table 10 Average PSNR and SSIM of dehazed results on the D-HAZE dataset [64]

	DCP	BCCR	NLD	DehazeNet	MSCNN	AOD-Net	DCPDN	HR-Dehazer	RI-GAN	Big FastNet	Dense 123	DRDDN
PSNR	14.56	14.62	13.18	11.36	12.52	12.80	12.04	16.47	16.47	16.37	19.92	**20.09**
SSIM	0.40	0.35	0.36	0.37	0.37	0.42	0.43	0.52	0.54	0.57	0.65	**0.67**

Bold indicates the best result

SSIM and PSNR—higher is better

respect to the PSNR score and SSIM. It should be pointed out that our model is trained on Dense-Haze training dataset. We display a dehazing example in Fig. 5, which shows that our method can generate a visually pleased dehazed results. *Evaluations on the D-Hazy-MB dataset* [65] To further show the generalization of the proposed dehazing method, we evaluate the dehazing performance compared with state-of-the-art methods on the cross-domain dataset D-Hazy-MB [65]. Table 11 lists the average PSNR and SSIM on D-Hazy-MB, where we show the best results by bold. DCP, CAP, DehazeNet and our method show comparable performance on SSIM, but our method obtains a higher PSNR, which indicates that our method can better restore haze-free images.

4.5 Qualitative comparison

In order to show the generalization ability of the proposed method in dealing with real-world natural hazy images, we test the proposed method on several natural challenging hazy images provided by previous methods [60] and collected from the Internet. The dehazed results are visually compared against six state-of-the-art methods. The results for six sample images obtained from the previous methods [9,12–14,18,70] are shown in Figs. 6 and 7. As shown in Figs. 6 and 7, we can notice that the results of BCCR [9] and CAP [12] tend to produce color distortion, while the results of methods [13,14,18] tend to retain haze. GFN suffers from the problem of color distortion. As shown in Fig. 6, PCFAN [70] cannot deal dense haze images well. Our model is able to capture the haze distribution well due to the larger receptive of network, which is critical to remove non-uniform haze. Low-level features are critical to recover the texture and appearance of haze-free image, and the model propagates the low-level features to deep layers via dense residual learning. So our method is able to generate realistic colors while removing haze moderately.

Furthermore, we show the high performance by comparing with four end-to-end methods [18–21]. As shown in Fig. 8, AOD-Net tends to leave haze in dehazed result, which makes the dehazed result lose image details. GFN tends to produce color distortion. FEED-Net tends to produce some noise in dehazed result, which is obvious in the zoomed-in "lake" image. DCPDN cannot keep the details in dehazed result well, which can be seen in "reiverside" image. Our model propagates the low-level features to deep layers and uses them to recover the final dehazed image, which is critical to preserve the image details. Comparing with other deep learning-based dehazing method, our method can remove haze well without losing the image details (Table 12).

In order to further show the performance of dehazing methods, we evaluate the dehazing using the metrics proposed in [71]. The value of e evaluates the ability of the dehazing method to restore edges, the value of $\bar{r}$ expresses

Table 11 Average PSNR/SSIM of dehazed results on hazy images from **D-Hazy-MB** [65]

	DCP	BCCR	CAP	NLD	MSCNN	DehazeNet	AOD-Net	DCPDN	GFN	DRDDN
PSNR	13.86	15.09	13.90	14.40	13.70	13.86	15.05	12.54	13.32	**16.43**
SSIM	0.77	0.73	**0.78**	0.70	0.74	0.77	0.59	0.71	0.76	0.77

(a) Inputs (b) DCP (c) NLD (d) Big FastNet (e) 123-CEDH (f) DRDDN (g) Ground Truths
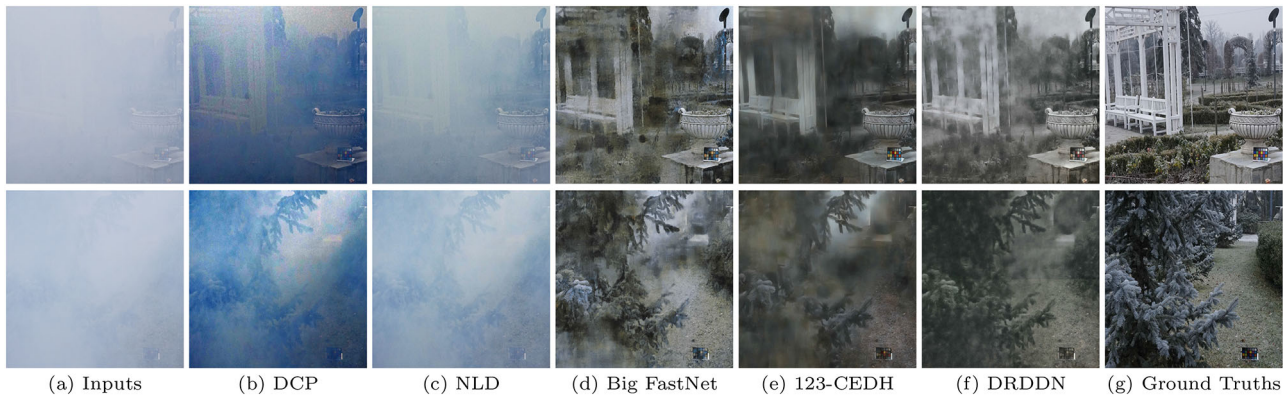
Fig. 5 Qualitative evaluations on the dense hazy images. The proposed method obtains much clearer images with clearer structures and characters

the quality of the contrast restoration, while the value of σ indicates the saturating of restored image. Table 14 shows the performance of dehazing methods. Higher the values of e show that the ability of restoration edges of dehazing method is more powerful. Higher the values of $\bar{r}$ indicate the better contrast of dehazed image. The value of σ shows the saturation of dehazed image. Tables 13 and 14 list the metrics of dehazing methods on natural images. From Table 13, we can see that our method can recover more powerful edges than other methods and avoid the over saturation problem. From Table 14, we can see the proposed method has more texture information and significant edges than the existing haze removal methods, recovers contrast better than other methods and avoids the over saturation problem.

4.6 Limitations

It shall point out that there exists a gap between different datasets, which incurs the degradation of the proposed model. Each dataset can be treated as domain and deep networks can be weak at generalizing learned knowledge to new datasets or environments [72]. Even a subtle change from the training domain can cause deep networks to make spurious predictions on the target domain. Such a problem is also true for dehazing models. The reason is twofold. First, a dehazing model may not work well when testing samples are different from synthesized training images. Second, results could be unreliable if the expected output is far from what the model is trained. We will address these problems by dedicating to collect more comprehensive clean samples and simulated more training samples.

5 Conclusion

In this paper, we present a deep fully end-to-end dehazing network. Our method is composed of residual learning and dilated densely connected block. The proposed model, termed as DRDDN, maximizes the information between layers, which makes our model learn more effective features for image dehazing. Extensive experiments are performed on three synthetic hazy datasets and real-world hazy images, which show the promising performance of the proposed method as compared to the recent state-of-the-art methods. In addition, an ablation study is conducted to show the improvements obtained by the incorporation of deep residual learning and dilated densely connected block to the recent state-of-the-art methods. In addition, an ablation study is performed to demonstrate the improvements obtained by the incorporation of deep residual learning and dilated densely connected block.
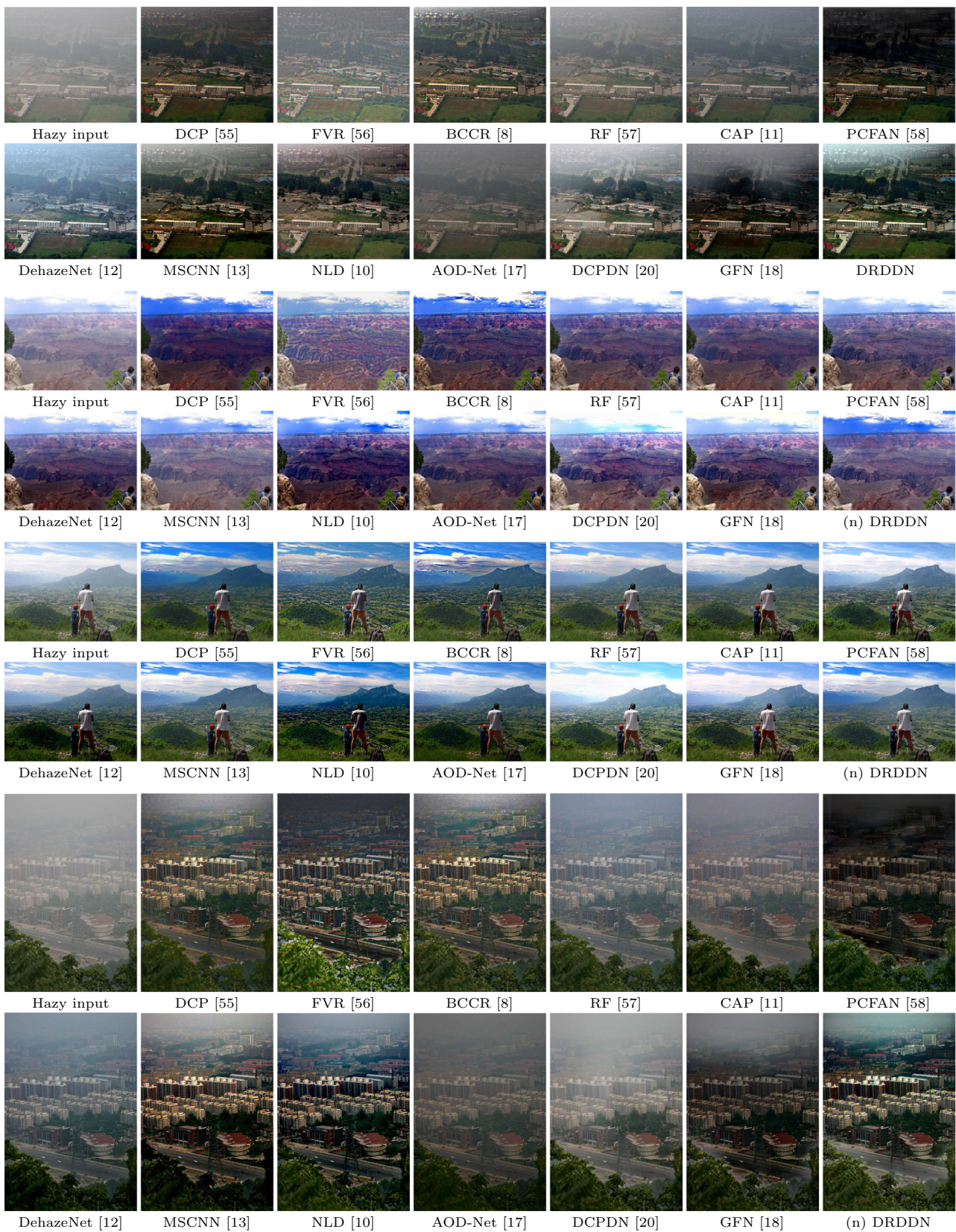
Fig. 6 The dehazing result of the proposed method and state-of-the-art methods on the real hazy images. The proposed method generates much clearer images with clearer structures and characters
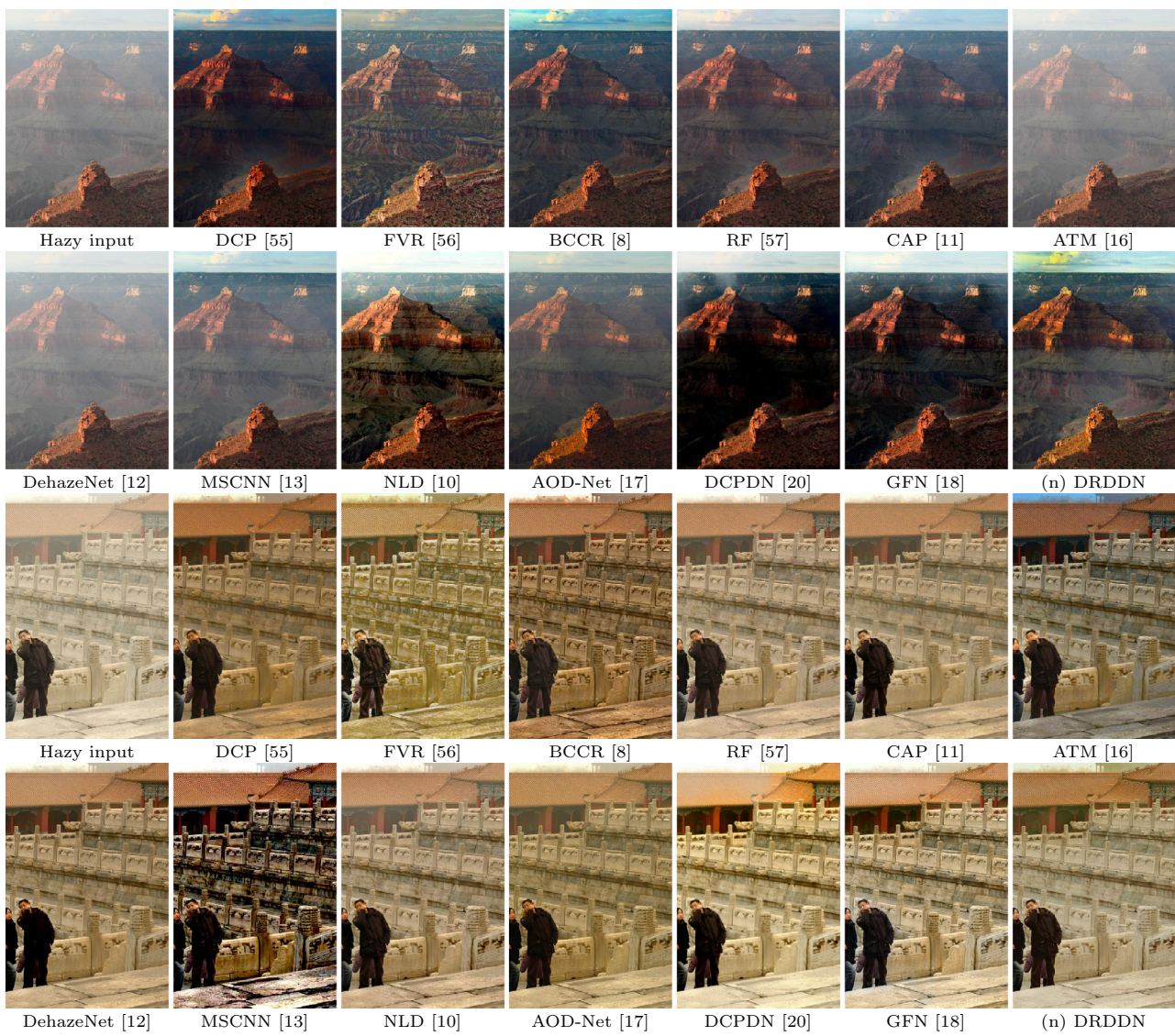
Fig. 7 The dehazing result of the proposed method and state-of-the-art methods on real-world images

Fig. 8 The dehazing result of the proposed method and state-of-the-art methods on real-world dense hazy images. In the result of FEED-Net, there are artifacts shown in blue box. AOD-Net tends to leave haze in the results. Note the lacking details in third row result of DCPCN. Note the color distortion in result of GFN in first row. Our method removes haze moderately and preserves the image details well

Table 12 Configurations of the dilation block (DB)

Layer	conv1	conv2	conv3	conv4	conv5
Size	3	3	3	3	3
Channel concat	16	16	16	16	16
Dilation rate	8	8	8	8	8

DB is used to replace the DDCB, and the dilation rate is determined by the dilation rate in the proposed network

Table 13 Quantitative evaluation on the first image in Fig. 7 using the metrics proposed in [71]

	DCP	BCCR	CAP	NLD	MSCNN	DehazeNet	AOD-Net	DCPDN	GFN	DRDDN
e	1.62	2.18	0.92	1.87	1.39	2.05	0.96	1.80	2.10	2.26
σ	0.08	0.05	0.03	9.70	0.00	7.12	0.31	19.32	1.35	0.05
$\bar{r}$	4.64	1.78	1.27	4.05	1.87	1.26	1.25	1.78	2.56	2.46

High values of e and $\bar{r}$ and low values of σ show the better result

Table 14 Quantitative evaluation on the first image in Fig. 8 using the metrics proposed in [71]

	AOD-Net	FEED-Net	DCPDN	GFN	DRDDN
e	2.95	36.21	3.26	25.8179	36.29
σ	0.21	0.27	0.23	2.08	0.11
$\bar{r}$	2.46	4.29	3.60	3.59	5.59

High values of e and $\bar{r}$ and low values of σ show the better result

Declaration

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Fan, C., Peng, Y., Peng, S., Zhang, H., Wu, Y., Kwong, S.: Detection of train driver fatigue and distraction based on forehead EEG: a time-series ensemble learning method. IEEE Trans. Intell. Transport. Syst. (2021). https://doi.org/10.1109/TITS.2021.3125737
2. Li, Q., Li, L., Wang, W., Li, Q., Zhong, J.: A comprehensive exploration of semantic relation extraction via pre-trained CNNs. Knowl. Based Syst. **194**, 105488 (2020)
3. Saini, M.K., Wang, X., Atrey, P.K., Kankanhalli, M.S.: Adaptive workload equalization in multi-camera surveillance systems. IEEE Trans. Multimed. **14**(3–1), 555–562 (2012)
4. Narasimhan, S.G., Nayar, S.K.: Vision and the atmosphere. Int. J. Comput. Vis. **48**(3), 233–254 (2002)
5. Schechner, Y.Y., Narasimhan, S.G., Nayar, S.K.: Instant dehazing of images using polarization. In: IEEE Conference on Computer Vision and Pattern Recognition (2001)
6. Narasimhan, S.G., Nayar, S.K.: Contrast restoration of weather degraded images. IEEE Trans. Pattern Anal. Mach. Intell. **25**(6), 713–724 (2003)
7. Shwartz, S., Namer, E., Schechner, Y.Y.: Blind haze separation. In: 2006 IEEE on Computer Vision and Pattern Recognition, vol. 2, pp. 1984–1991. IEEE (2006)

8. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. In: IEEE Conference on Computer Vision and Pattern Recognition (2009)
9. Meng, G., Wang, Y., Duan, J., Xiang, S., Pan, C.: Efficient image dehazing with boundary constraint and contextual regularization. In: IEEE International Conference on Computer Vision (2013)
10. Fattal, R.: Dehazing using color-lines. ACM Trans. Graph. **34**(1), 13 (2014)
11. Berman, D., Treibitz, T., Avidan, S.: Non-local image dehazing. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1674–1682 (2016)
12. Zhu, Q., Mai, J., Shao, L.: A fast single image haze removal algorithm using color attenuation prior. IEEE Trans. Image Process. **24**(11), 3522–3533 (2015)
13. Cai, B., Xu, X., Jia, K., Qing, C., Tao, D.: Dehazenet: an end-to-end system for single image haze removal. IEEE Trans. Image Process. **25**(11), 5187–5198 (2016)
14. Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., Yang, M.-H.: Single image dehazing via multi-scale convolutional neural networks. In: European Conference on Computer Vision (2016)
15. Ren, W., Pan, J., Zhang, H., Cao, X., Yang, M.-H.: Single image dehazing via multi-scale convolutional neural networks with holistic edges. Int. J. Comput. Vis. **128**, 1–20 (2019)
16. Berman, D., Treibitz, T., Avidan, S.: Air-light estimation using haze-lines. In: 2017 IEEE International Conference on Computational Photography (ICCP), pp. 1–9. IEEE (2017)
17. Sulami, M., Glatzer, I., Fattal, R., Werman, M.: Automatic recovery of the atmospheric light in hazy images. In: IEEE International Conference on Computational Photography (2014)
18. Li, B., Peng, X., Wang, Z., Xu, J., Feng, D.: An all-in-one network for dehazing and beyond. In: IEEE International Conference on Computer Vision (2017)
19. Ren, W., Ma, L., Zhang, J., Pan, J., Cao, X., Liu, W., Yang, M.-H.: Gated fusion network for single image dehazing. In: IEEE Conference on Computer Vision and Pattern Recognition (2018)
20. Zhang, S., Ren, W., Yao, J.: Fully end-to-end dehazing. In: ICME, Feed-net (2018)
21. Zhang, H., Patel, V.M.: Densely connected pyramid dehazing network. In: IEEE Conference on Computer Vision and Pattern Recognition (2018)
22. Zhang, S., He, F.: DRCDN: learning deep residual convolutional dehazing networks. Vis. Comput. **36**, 1–12 (2019)
23. Yu, F., Koltun, V., Funkhouser, T.A.: Dilated residual networks. In: IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, p. 3 (2017)
24. Zhu, Y., Tang, G., Zhang, X., Jiang, J., Tian, Q.: Haze removal method for natural restoration of images with sky. Neurocomputing **275**, 499–510 (2018)
25. Zhang, S., Yao, J., Garcia, E.B.: Single image dehazing via image generating. In: Pacific-Rim Symposium on Image and Video Technology, pp. 123–136. Springer, Berlin (2017)
26. Zhang, S., Yao, J.: Single image dehazing using fixed points and nearest-neighbor regularization. In: Asian Conference on Computer Vision, pp. 18–33 (2016)

27. Zhang, S., He, F., Yao, J.: Single image dehazing using deep convolution neural networks. In: Pacific Rim Conference on Multimedia, pp. 315–325. Springer, Berlin (2017)

28. Zhang, S., He, F., Ren, W., Yao, J.: Joint learning of image detail and transmission map for single image dehazing. Vis. Comput. **36**(2), 305–316 (2020)

29. Kim, J., Lee, J.K., Lee, K.M.: Accurate image super-resolution using very deep convolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1646–1654 (2016)

30. Ren, W., Zhang, J., Ma, L., Pan, J., Cao, X., Zuo, W., Liu, W., Yang, M.-H.: Deep non-blind deconvolution via generalized low-rank approximation. In: Advances in Neural Information Processing Systems, vol. 18, pp. 297–307 (2018)

31. Yan, Y., Ren, W., Cao, X.: Recolored image detection via a deep discriminative model. IEEE Trans. Inf. Forensics Secur. **14**(1), 5–17 (2018)

32. Ren, W., Liu, S., Ma, L., Xu, Q., Xu, X., Cao, X., Du, J., Yang, M.-H.: Low-light image enhancement via a deep hybrid network. IEEE Trans. Image Process. **28**(9), 4364–4375 (2019)

33. Ren, W., Zhang, J., Pan, J., Liu, S., Ren, J., Du, J., Cao, X., Yang, M.-H.: Deblurring dynamic scenes via spatially varying recurrent neural networks. IEEE Trans. Pattern Anal. Mach. Intell. (2021). https://doi.org/10.1109/TPAMI.2021.3061604

34. He, Z., Cao, Y., Du, L., Xu, B., Zhuang, Y.: MRFN: multi-receptive-field network for fast and accurate single image super-resolution. IEEE Trans. Multimed. **PP**(99), 1 (2019)

35. Tan, X., Zhu, H., Shao, Z., Hou, X., Hao, Y., Ma, L.: Saliency detection by deep network with boundary refinement and global context. In: 2018 IEEE International Conference on Multimedia and Expo (ICME), pp. 1–6. IEEE (2018)

36. Tan, X., Xu, K., Cao, Y., Zhang, Y., Ma, L., Lau, R.W.H.: Night-time scene parsing with a large real dataset. IEEE Trans. Image Process. **30**, 9085–9098 (2021)

37. Ren, W., Zhang, J., Xu, X., Ma, L., Cao, X., Meng, G., Liu, W.: Deep video dehazing with semantic segmentation. IEEE Trans. Image Process. **28**(4), 1895–1908 (2019)

38. Zhang, S., He, F., Ren, W.: NLDN: non-local dehazing network for dense haze removal. Neurocomputing **410**, 363–373 (2020)

39. Li, L., Dong, Y., Ren, W., Pan, J., Gao, C., Sang, N., Yang, M.-H.: Semi-supervised image dehazing. IEEE Trans. Image Process. **29**, 2766–2779 (2019)

40. Ren, W., Pan, J., Zhang, H., Cao, X., Yang, M.-H.: Single image dehazing via multi-scale convolutional neural networks with holistic edges. Int. J. Comput. Vis. **128**(1), 240–259 (2020)

41. Zhang, S., Ren, W., Tan, X., Wang, Z.-J., Liu, Y., Zhang, J., Zhang, X., Cao, X.: Semantic-aware dehazing network with adaptive feature fusion. IEEE Trans. Cybern. (2021). https://doi.org/10.1109/TCYB.2021.3124231

42. Li, S., Ren, W., Wang, F., Araujo, I.B., Tokuda, E.K., Junior, R.H., Cesar-Jr, R.M., Wang, Z., Cao, X.: A comprehensive benchmark analysis of single image deraining: current challenges and future perspectives. Int. J. Comput. Vis. **129**(4), 1301–1322 (2021)

43. Liu, X., Ma, Y., Shi, Z., Chen, J.: Griddehazenet: attention-based multi-scale network for image dehazing. In: IEEE International Conference on Computer Vision (2019)

44. Qin, X., Wang, Z., Bai, Y., Xie, X., Jia, H.: FFA-NET: feature fusion attention network for single image dehazing. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, pp. 11908–11915 (2020)

45. Chen, Z., Wang, Y., Yang, Y., Liu, D.: PSD: principled synthetic-to-real dehazing guided by physical priors. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7180–7189 (2021)

46. Zheng, Z., Ren, W., Cao, X., Hu, X., Wang, T., Song, F., Jia, X.: Ultra-high-definition image dehazing via multi-guided bilat-eral learning. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 16185–16194 (2021)

47. Zhang, J., Ren, W., Zhang, S., Zhang, H., Nie, Y., Xue, Z., Cao, X.: Hierarchical density-aware dehazing network. IEEE Trans. Cybern. (2021). https://doi.org/10.1109/TCYB.2021.3070310

48. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)

49. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, p. 3 (2017)

50. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: ICLR (2015)

51. Chen, L.-C., Papandreou, G., Schroff, F., Adam, H.: Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587 (2017)

52. Yang, D., Sun, J.: Proximal dehaze-net: a prior learning-based deep network for single image dehazing. In: ECCV, pp. 702–717 (2018)

53. Silberman, N., Hoiem, D., Kohli, P., Fergus, R.: Indoor segmentation and support inference from RGBD images. In: European Conference on Computer Vision (2012)

54. Liu, F., Shen, C., Lin, G., Reid, I.: Learning depth from single monocular images using deep convolutional neural fields. IEEE Trans. Pattern Anal. Mach. Intell. **38**(10), 2024–2039 (2016)

55. Yu, F., Koltun, V.: Multi-scale context aggregation by dilated convolutions. arXiv preprint arXiv:1511.07122 (2015)

56. Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W., Wang, Z.: Benchmarking single image dehazing and beyond. TIP **28**, 492–505 (2018)

57. Fahim, M.A.-N.I., Jung, H.Y.: Single image dehazing using end-to-end deep-dehaze network. Electronics **10**(7), 817 (2021)

58. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. IEEE Trans. Pattern Anal. Mach. Intell. **33**(12), 2341–2353 (2011)

59. Tarel, J.-P., Hautiere, N.: Fast visibility restoration from a single color or gray level image. In: IEEE International Conference on Computer Vision (2009)

60. Tang, K., Yang, J., Wang, J.: Investigating haze-relevant features in a learning framework for image dehazing. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2995–3000 (2014)

61. Chen, C., Do, M.N., Wang, J.: Robust image and video dehazing with visual artifact suppression via gradient residual minimization. In: European Conference on Computer Vision (2016)

62. Zhang, Y., Ding, L., Sharma, G.: Hazerd: an outdoor scene dataset and benchmark for single image dehazing. In: ICIP, pp. 3205–3209. IEEE (2017)

63. Ancuti, C.O., Ancuti, C., Timofte, R., De Vleeschouwer, C.: O-haze: a dehazing benchmark with real hazy and haze-free outdoor images. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 754–762 (2018)

64. Ancuti, C.O., Ancuti, C., Sbert, M., Timofte, R.: Dense haze: a benchmark for image dehazing with dense-haze and haze-free images. arXiv preprint arXiv:1904.02904 (2019)

65. Ancuti, C., Ancuti, C.O., De Vleeschouwer, C.: D-hazy: a dataset to evaluate quantitatively dehazing algorithms. In: IEEE International Conference on Image Processing, pp. 2226–2230. IEEE (2016)

66. Morales, P., Klinghoffer, T., Lee, S.J.: Feature forwarding for efficient single image dehazing. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (2019)

67. Dudhane, A., Singh Aulakh, H., Murala, S.: Ri-gan: an end-to-end network for single image haze removal. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (2019)

68. Bianco, S., Celona, L., Piccoli, F., Schettini, R.: High-resolution single image dehazing using encoder–decoder architecture. In:

IEEE Conference on Computer Vision and Pattern Recognition Workshops (2019)

69. Guo, T., Cherukuri, V., Monga, V.: Dense123'color enhancement dehazing network. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops (2019)

70. Zhang, X., Wang, T., Wang, J., Tang, G., Zhao, L.: Pyramid channel-based feature attention network for image dehazing. Comput. Vis. Image Underst. **197**, 103003 (2020)

71. Hautière, N., Tarel, J.-P., Aubert, D., Dumont, E.: Blind contrast enhancement assessment by gradient ratioing at visible edges. Image Anal. Stereol. **27**(2), 87–95 (2008)

72. Tzeng, E., Hoffman, J., Darrell, T., Saenko, K.: Simultaneous deep transfer across domains and tasks. In: IEEE International Conference on Computer Vision, pp. 4068–4076 (2015)

Fazhi He received Ph.D. degree from Wuhan University of Technology. He was post-doctor researcher in The State Key Laboratory of CAD&CG at Zhejiang University, a visiting researcher in Korea Advanced Institute of Science & Technology and a visiting faculty member in the University of North Carolina at Chapel Hill. Now, he is a professor in School of Computer, Wuhan University. His research interests are computer graphics, computer-aided design, image processing, and computer supported cooperative work.
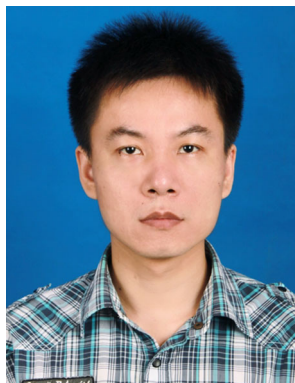
Neng Hou received Ph.D degree from Wuhan University in 2018. He is currently a lecture at School of Computer Science in Yangtze University. His research interests include global optimization, GPU computing, and deep learning.

Shengdong Zhang is currently a Ph.D candidate at the School of Computer Science in Wuhan University. He received the B.S. degree in computer science from Shanxi University, China, in 2009, and the M.S. degree in mechanics from Peking University, in 2012. His research interests include deep leaning and image processing.

Jiaoting Zhang received her M.S. degree from Civil Aviation University of China, Tianjin, China, in 2018. Her research interests are pattern recognition, image processing, and computer graphics.