IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 2 July 2023, accepted 9 July 2023, date of publication 18 July 2023, date of current version 15 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3296592

RESEARCH ARTICLE

Dual Multi-Scale Dehazing Network

SHENGDONG ZHANG[1,2], XIAOQIN ZHANG[2], AND LINLIN SHEN[1], (Senior Member, IEEE)

[1]College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China
[2]Key Laboratory of Intelligent Informatics for Safety and Emergency of Zhejiang Province, Wenzhou University, Wenzhou 325035, China

Corresponding authors: Xiaoqin Zhang (zhangxiaoqinnan@gmail.com) and Linlin Shen (llshen@szu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant U2033210, in part by the Zhejiang Provincial Natural Science Foundation under Grant LDT23F02024F02, and in part by the Science and Technology Plan Project in Basic Public Welfare Class of Shaoxing City under Grant 2022A11002.

ABSTRACT Single-image haze removal is a challenging ill-posed problem. Recently, methods based on training on synthetic data have achieved good dehazing results. However, we note that these methods can be further improved. A novel deep learning-based method is proposed to obtain a better-dehazed result for single-image dehazing in this paper. Specially, we propose a dual multi-scale network to learn the dehazing knowledge from synthetical data. The coarse multi-scale network is designed to capture a large variety of objects, and then fine multi-scale blocks are designed to capture a small variety of objects at each scale. To show the effectiveness of the proposed method, we perform experiments on a synthetic dataset and real hazy images. Extensive experimental results show that the proposed method outperforms the state-of-the-art methods.

INDEX TERMS Dual multi-scale, dehazing, synthetically data, deep learning, real haze image.

I. INTRODUCTION

The turbid medium in the atmosphere often degrades the image quality. Outdoor images taken in bad weather tend to show a hazy and blurry appearance. Atmospheric absorption and scattering cause haze, which reduces the contrast and fades the color of outdoor images. The light reaches by the camera from the scene objects is attenuated along the line of sight and blended with the atmospheric light. The absorption and scattering processes are commonly modeled by a linear combination of the direct attenuation and the air-light [1]:

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

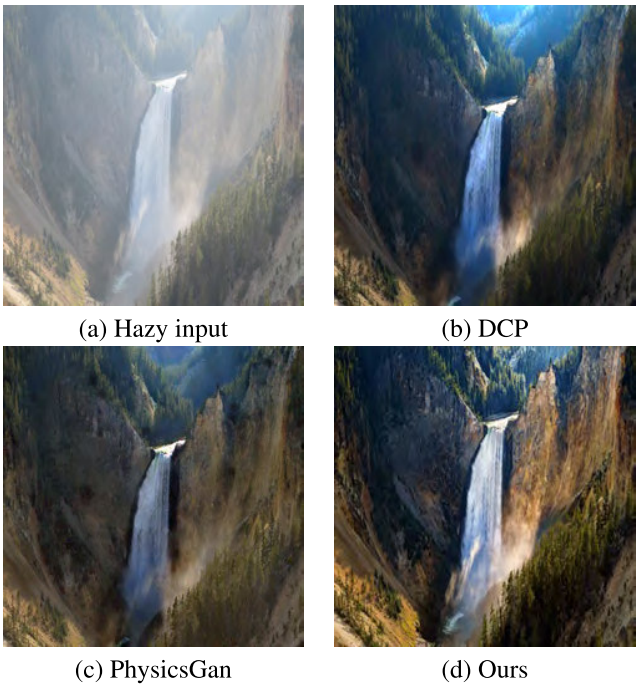
where the I is the input hazy image, and the J it the corresponding clean image, t represents how much the light reflected from objects is received by camera, A is the air-light.

Single-image dehazing, which aims at removing haze from single input image as much as possible, has a wide variety of applications, such as auto driving, semantic segmentation, image recognition, etc. Due to its wide applications, dehazing has attracted much attention. There are two key steps in the dehazing process: 1) estimation of transmission map and atmospheric light, and 2) compute the final dehazed result. Prior-based methods [2], [3] have been proposed to remove haze based on two steps of dehazing. Due to the fact prior is based on simple statistic law, which cannot be satisfied by real cases. For example, dark channel prior (DCP) [2] cannot deal the white objects well.

Inspired by the success of data-driven methods, many researchers proposed end-to-end CNN models [4], [5], [6], [7], [8] for single-image dehazing. Although these methods have shown effectiveness on a synthetic dataset. However, these methods have limitations due to large-scale arbitrariness caused by haze. Furthermore, The distribution of haze is depend on depth, which needs different receptive field sizes to estimate the depth for each pixel.

To overcome these two issues jointly, we propose a dual multi-scale dehazing network. The formation of haze can be affected by various factors, such as temperature, altitude, and humidity, making the distribution of haze at individual spatial locations space-variant and non-homogeneous. To capture the distribution of haze, we propose a dual multi-scale dehazing network, which has different perceptive fields and captures objects with different sizes. We compare our method with traditional and learning-based methods [2], [9] in Fig 1. The main contributions of this work are listed as follows:

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



(a) Hazy input (b) DCP

(c) PhysicsGan (d) Ours

FIGURE 1. Visual comparisons on a challenging real-world hazy example. The dehazed result of DCP tends to show a dark appearance and the tree area cannot be recovered well, compared with (d). The dehazed result of PhysicsGan looks better than DCP, there is still room for improvement. Compared with DCP and PhysicsGan, our method often generates a visual favorable result.

- We propose a dual multi-scale dehazing network, which can capture the large and small variety of objects and understand the distribution of haze. The distribution of haze is very large, we employ a coarse multi-scale network to capture the global haze distribution. We then capture the small variety of distribution of haze via fine multi-scale blocks. The proposed model can capture the global and local distribution of haze well and effectively improves the dehazing performance.
- We propose a fine multi-scale block, which can capture small varieties of objects. The distribution of haze depends on the depth, which is different for different objects. However, the distribution of haze within one object tends to show homogeneous. It is critical to design a network that can capture small varieties of object sizes in each scale, which motivates us to design a fine multi-scale block.
- We conduct extensive experiments to quantitatively and qualitatively compare the proposed method with the state-of-the-art single-image dehazing methods and demonstrate the effectiveness of the proposed model.

II. RELATED WORK

Single image dehazing methods can be mainly grouped into two approaches: physical model based recovering methods and color information based enhancement approaches.

A. SINGLE IMAGE DEHAZING METHODS

Physical models based methods [10], [11] assume that hazy images can be modeled by Eq. (1), which models hazy images as the linear sum of clean image and atmospheric light [12]. Clean image means the scene information that are not affected by medium particles. Based on this model, most existing algorithms focus on recovering the scene that not reaches the camera sensors, *i.e.,* estimating the transmission map $t(x)$ for each hazy image. For example, a improved image formation model is proposed by [13]. This model is designed for the estimating of transmission map and surface shading. The hazy image can be treat as regions of constant albedo, and we can infer the the scene transmission from hazy image. Dark channel prior (DCP) is inferred from the features of non-sky haze-free images. The DCP assumes the at least one pixel contains a channel whose intensity is close to zero. [10] extends the DCP and proposes a more general boundary constraint. Four haze-relevant priors are studied [14] and a multi-scale dehazing method are designed to improve the dehazing performance. Reference [15] finds that the relation between brightness and saturation in a clear image patch, and proposes a color attenuation prior to compute transmission maps. Reference [3] finds that a clean image can be presented by hundreds of color clusters. However, a hazy image cannot be presented by hundreds of color clusters. Based on this observation, [3] design a non-local method to compute the transmission map. However, these hand-crafted priors are statistical properties over a large number of images and thus cannot hold always in practical scenarios. For example, when the scene objects are close to the airlight, the dark channel has bright values near such objects, which means that the dark channel prior is not hold, and as a result the haze layer will be overestimated [2].

To avoid designing statistical features, several algorithms employ deep convolutional neural networks (CNN) to improve image dehazing. Both DehazeNet [11] and MSCNN [16] use a deep neural network for transmission estimation and then follow the conventional method to estimate atmospheric light and haze-free image. Instead of computing the transmission map and the atmospheric light separately, AOD-Net [17] incorporates the transmission and the airlight into a new variable and design an light dehazing method. However, this method tends to retain haze in dehazed result. DCPDN [18] and DDN [19] are two methods, which incorporate the scattering model into deep network. These methods need two networks to compute transmission maps and atmospheric lights first, then restore final dehazed images by inversing the model (1). An end-to-end fusion-based dehazing network [20] is proposed to predict weight maps to combine three derived inputs into a single one by choosing the most important features of them. However, GFN also computes three inputs using traditional methods and intermediate confidence maps were needed to be computed. Qin et al. design a novel pixel and channel attention [5] to improve the dehazing performance. Pan et al. design a physics-based

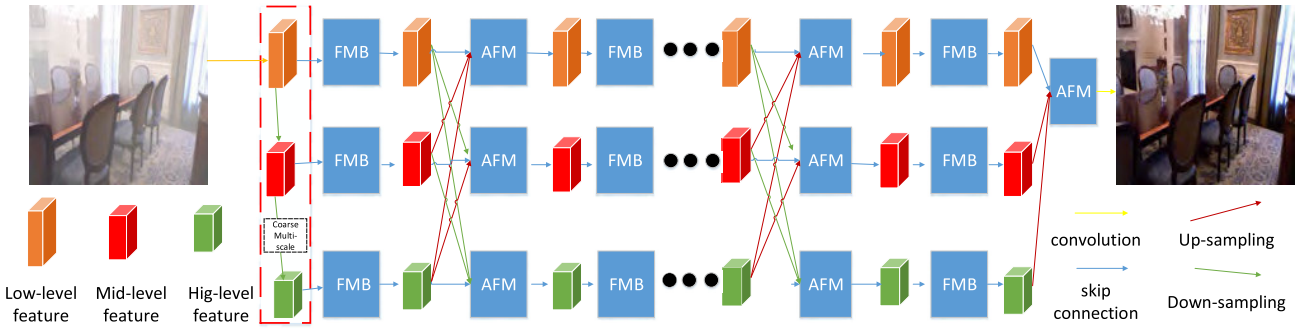


FIGURE 2. The architecture of the proposed network. The coarse multi-scale network contains three different scales, which are used to capture global and local information. Another core of the proposed network is the fine multi-scale block, which is designed to choose the most important feature from multi-scale features. The adaptive fusion module is designed to fuse the multi-scale features from the coarse multi-scale network.

generative networks [9] for image restoration problem, which can incorporate the physics model to boost the dehazing performance. Dong et al. employ the boosting strategy to design a multi-scale dehazing network [21]. Zheng et al. study the ultra-high-definition image dehazing [22] based on the physical model. Although promising results have been obtained, the assumption that hazy images is the sum of clean image and airlight does not hold in real complex scenes, especially when the haze is heavy and contains noise. To improve the dehazing performance on natural hazy images, Shao et al. propose a domain adaptation dehazing method [23]. Different from these methods, our method takes multi-scale ability into the proposed network for dehazing and achieves the fast dehazing performance.

Prior based dehazing methods can restore dehazed sharp results at the expense of low quantitative results for synthetic images. Data-driven dehazing methods obtain high quantitative results for synthetic images but cannot remove haze from real hazy images completely. To address the disadvantages of prior based dehazing methods and data-driven dehazing methods, neural augmentation based dehazing methods [24], [25], [26] are proposed. Neural augmentation based dehazing methods estimate the atmospheric light and transmission map firstly, and then data-driven methods are used to refine the the atmospheric light and transmission map. The dehazed results are obtained by physical model with the estimated the atmospheric light and transmission map.

III. PROPOSED METHOD

The proposed model is a dual multi-scale dehazing network, the overall framework is shown in Fig. 2. The dual multi-scale ability is from coarse multi-scale network and fine multi-scale blocks. We firstly introduce the motivation, and then the dual multi-scale dehazing network, which learns dehazing ability from synthetic images.

A. MOTIVATION

Objects often have different sizes, which are hard for dehazing. As shown in Fig. 3, persons in red rectangles have

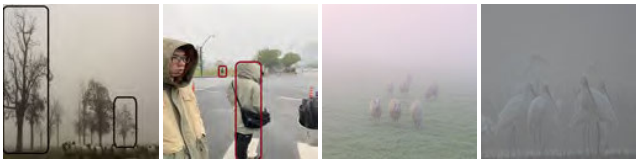


FIGURE 3. Visual examples for objects with different scales. The objects tend to show large variation in (a) and (b), while show small variation in (c) and (d).

different sizes due to the different depths, which result in different densities of haze for these areas. We note that the trees in black rectangles also have different sizes. We can see that the objects in near areas have large size, while the objects in far areas have small sizes. To capture such a dramatic variation of object sizes, we propose a coarse multi-scale network that increases the receptive field via down-sampling. As shown in Fig. 3(c) and (d), we note that the sheep and cranes have similar object sizes and show small variations of object sizes, it is important to capture such a variation for image dehazing. To capture such small variations of object sizes, we propose a fine multi-scale block, which employs different dilation rates to understand the variations in local areas. In order to capture the large and small variations of objects, we combine the coarse multi-scale network with fine multi-scale blocks.

B. DUAL MULTI-SCALE DEHAZING NETWORK

Based on the analysis in Section III-A, we propose a dual multi-scale dehazing network (DMSDN), the network detail can be found in Fig. 2. The dual multi-scale dehazing network consists of a coarse multi-scale network and fine multi-scale blocks. The coarse multi-scale network contains three scales. The first scale (coarse scale) contains six fine multi-scale blocks, the second (median scale) contains six fine multi-scale blocks, and the third scale contains six fine multi-scale blocks. To capture the global and local features, the model employs three scales of information to explore useful features for dehazing.

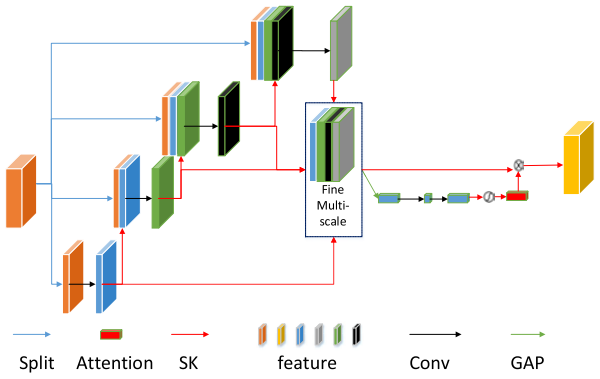


FIGURE 4. The architecture of the proposed fine multi-scale block (FMB). The block contains a dense dilated block and a channel-wise attention block. The dense dilated block is designed to extract multi-scale information, and then the channel-wise attention block chooses the most important information from the multi-scale information.

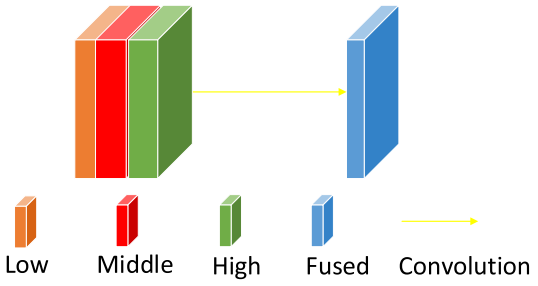


FIGURE 5. The architecture of the proposed adaptive fusion module. We first concat the multi-scale features and then use a convolution operation to extract effective feature for dehazing.

As the learned feature map exists redundant information, which is a reason for deep learning models cannot learn effective features for dehazing. In order to boost the learning efficient, we propose a fine multi-scale block. The fine multi-scale block (FMB) contains multi-scale information extracting and an attention module, which is shown in Fig 4. Based on the observation, the feature map contains redundant information, applying a convolution on it cannot learning information as much as possible. We split the feature into four sub-features, which contain sub-information of the original feature. We apply a convolution on one sub-feature, and obtain a new feature (O_1). We concat the O_1 with another sub-feature and obtain a concated feature (C_1), then we apply a convolution on C_1 and obtain a new feature (O_2). We repeat this process, and obtain O_3 and O_4. We concat O_1, O_2, O_3 and O_4, We then apply a channel attention on the concated feature and obtain the active feature for dehazing. The proposed module reduce the computation time and model complexity.

To further improve the information flow, we propose an adaptive fusion module (AFM), which fuses the features from each scale of a coarse multi-scale network adaptively. As shown in Fig. 5, we first concat the high-level, middle-level, and low-level features, and then a convolution operation with 1×1 kernel is applied, which obtains the fused feature.

C. TRAINING LOSS

Let $\mathcal{F}$ denote the mapping function which is learned by the network, and Θ represents the parameters of the network. Let $\{I_i, i = 1, 2, \cdots, N\}$ and $\{J_i, i = 1, 2, \cdots, N\}$ denote the hazy input images and the corresponding clean ones, respectively. It has been widely acknowledged that L_2 loss tends to produce blurry dehazed results [18]. To solve this issue efficiently, we introduce a novel edge-preserving loss, which is composed of two different parts: L_1 loss and perceptual loss. $\mathcal{L}_1$ is defined as follow:

$$\mathcal{L}_1 = \frac{1}{N}\sum_{i=1}^{N} \left\| \mathcal{F}(I_i, \Theta) - J_i \right\|_1, \tag{2}$$

where N is the number of training pair data.

To eliminate the visual artifacts of dehazed images, we employ perceptual loss to train the model. The perceptual loss consists of Feature Reconstruction Loss and Style Reconstruction Loss. Instead of encouraging the pixels of the dehazing image $\tilde{J}$ to exactly match the pixels of the ground truth image, feature reconstruction loss encourages them to have similar feature representations. The perceptual loss can be defined as follow:

$$\mathcal{L}_{per} = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{K} \left\| \phi_j(\tilde{J}_i) - \phi_j(J_i) \right\|_2^2, \tag{3}$$

where ϕ presents the VGG-19 network, which is trained on ImageNet, N demotes the number of training samples, and j denotes the layer number. We select the layers 'conv1-2', 'conv2-2', 'conv3-2', 'conv4-2', and 'conv5-2' in the VGG-19 network to compute the feature reconstruction loss.

Our overall loss function is:

$$\mathcal{L}(\tilde{J}, J) = \mathcal{L}_1 + \lambda_2 \mathcal{L}_{per}. \tag{4}$$

where λ_2 controls the contribution of perceptual loss.

D. IMPLEMENTATION DETAILS AND DATASET

In the proposed model, we set 3×3 as the kernel size for all convolution layers except the ones in AFM. In our experiment, we set the scale number to 3. For each scale-aware attention module, we set the dilation rate to 1, 2, 4 and 8. All dilated layers are initialized using an identity initializer [27]. We set λ_1=0.01 and λ_2=0.01 in all the experiments. We use a leaky rectified linear unit (LReLU) as our activation function. We use 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 0.0005, respectively. During training, we decrease the learning rate decreases half for every 30 epochs. The network was trained for totally 100 epochs by Pytorch with an Nvidia GTX 2018Ti GPU. We train the proposed network on the SOTS dataset from RESIDE [28] as the state-of-the-art dehazing methods [5], [29].

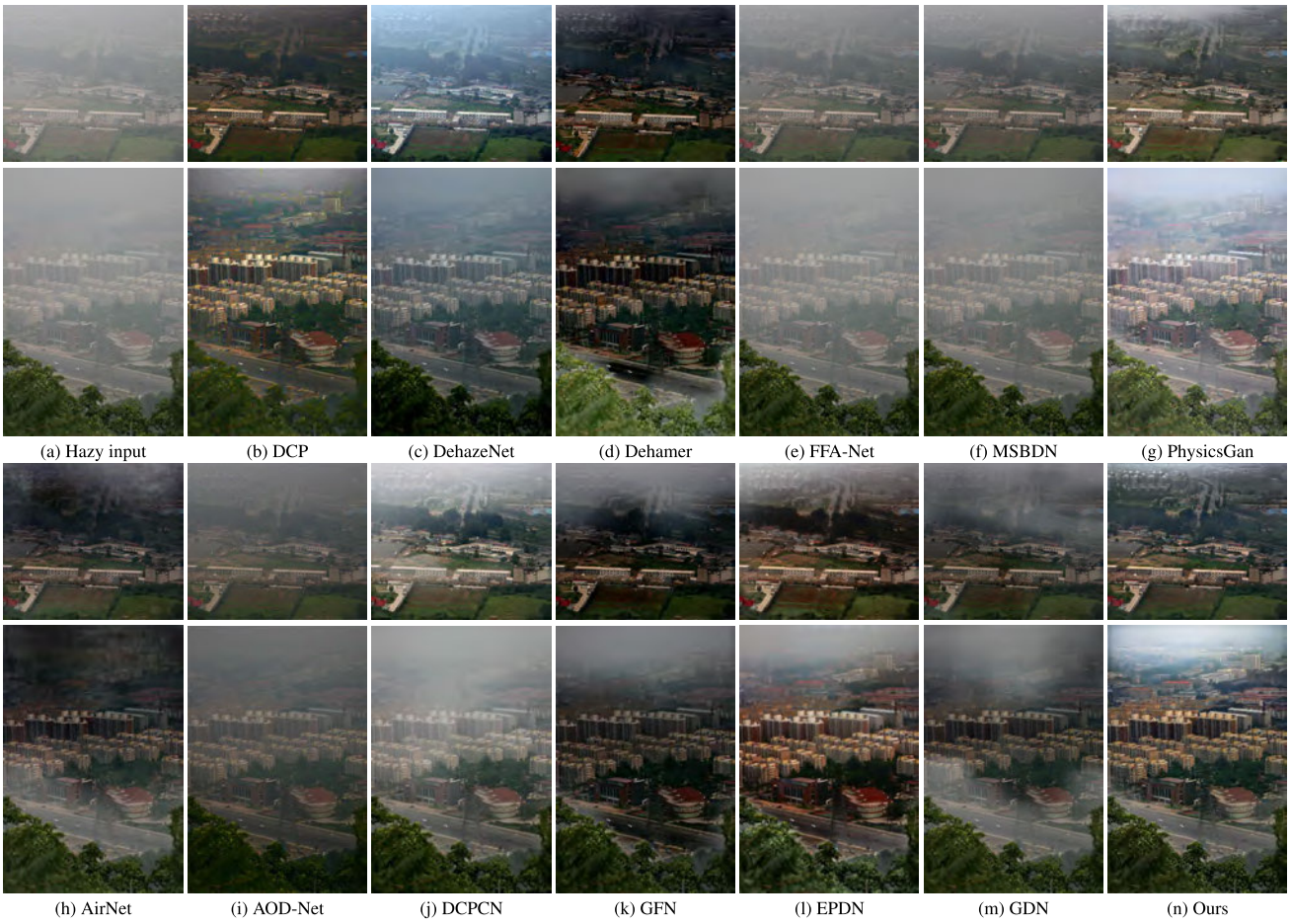


FIGURE 6. Visual comparisons of dehazed results of various dehazing methods and proposed method on real-world hazy images. It showed that dehazed results of the proposed method are much clearer than the results of other state-of-the-art methods.

TABLE 1. Average PSNR/SSIM of dehazed results on the SOTS dataset from RESIDE.

	DCP	NLD	MSCNN	DehazeNet	AOD-Net	GFN	GDN	EPDN	BPPNet	PhysicsGan	FFA-Net	MSBDN	AirNet	Ours
PSNR	16.62	17.29	17.57	21.14	19.06	22.30	32.16	25.06	27.16	29.43	**36.39**	33.79	20.58	34.80
SSIM	0.82	0.75	0.81	0.85	0.85	0.88	0.98	0.92	0.93	0.93	0.99	0.98	0.87	**0.99**

TABLE 2. Averaged DHQI scores for 79 natural hazy images. The best result is marked with red color, while the second best is marked with blue color.

DehazeNet	FFA-Net	MSDBN	EPDN	DCP	NLD	AOD-Net	Our
58.96	55.33	54.32	60.84	51.92	52.75	55.67	59.06

TABLE 3. Density values for a natural hazy image in Fig. 9. The best result is marked with red color, while the second best is marked with blue color.

Input	FFA-Net	AECR-Net	AirNet	EPDN	FAMEDNet	MSBDN	PSD-Net	SGID-PFF	Our
1.832	0.737	0.294	0.634	0.293	0.743	0.802	0.230	0.233	0.185

IV. EXPERIMENTAL RESULTS

Several state-of-the-art methods are compared in our experiment including Dark Channel Prior (DCP) [2], Non-Local Dehazing (NLD) [3], MSCNN [16], AOD-Net [17], Gated Fusion Network (GFN) [20], EPDN [30], GridDehazeNet [29], PhysicsGan [9], FFA-Net [5], MSBDN [21], AirNet [31], and Dehamer [32] on hazy images.

A. QUANTITATIVE COMPARISON

1) RESIDE DATASET

REalistic Single-Image DEhazing (RESIDE) [28] is the first large-scale simulated haze dataset, which provided indoor and outdoor hazy images. The hazy images in this dataset have the ground truth, we can evaluate the dehazing performance using PSNR and SSIM. The indoor part of

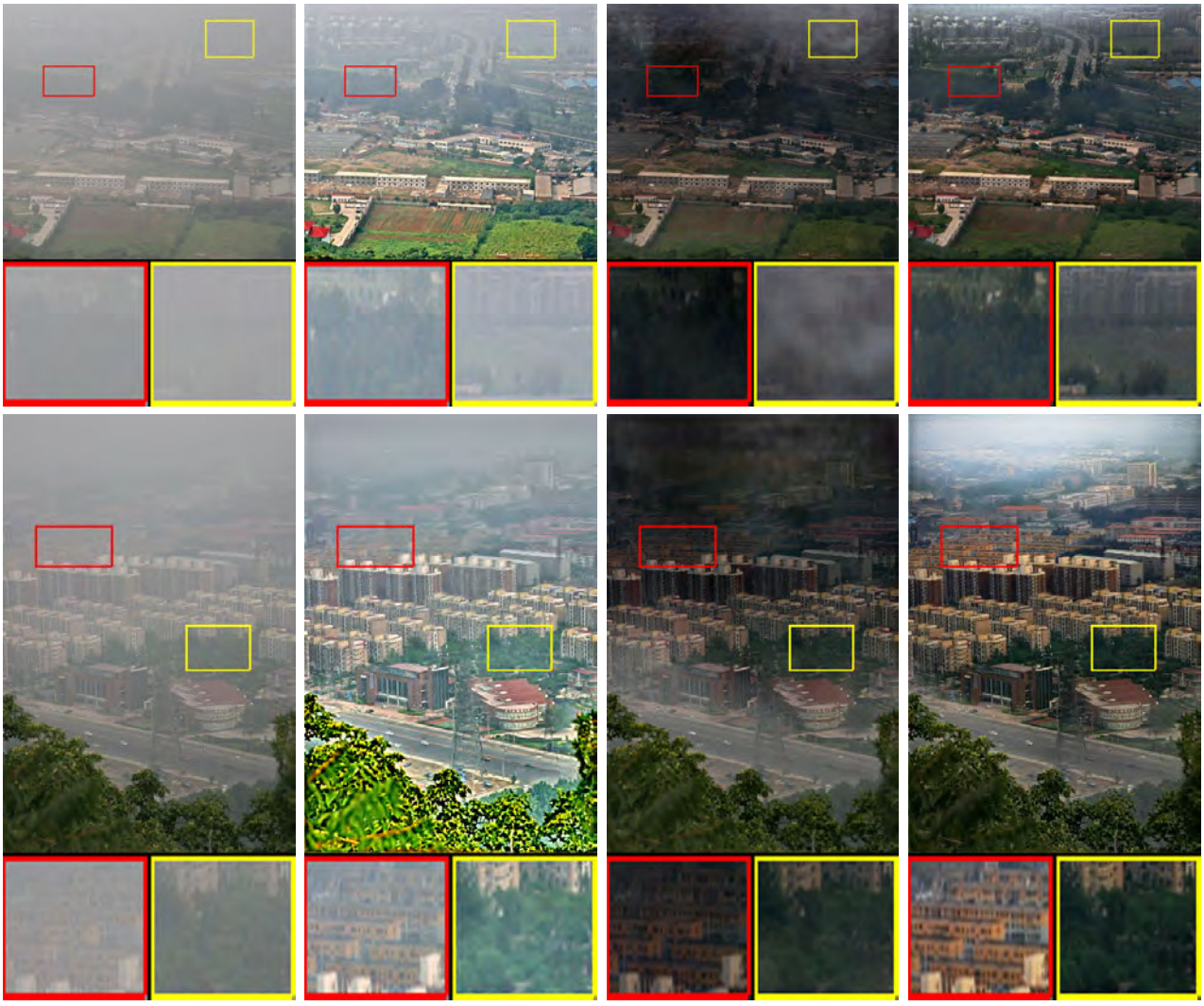


FIGURE 7. Visual results of some recently dehazing methods and proposed method. The PSD and AirNet employ hazy images to train the dehazing model. The dehazed results of PSD have a bright appearance and tend to leave haze. The dehazed results of AirNet have a dark appearance and tend to show haze. The proposed method tends to show a bright appearance and recover more image details and show a colorful dehazed results.

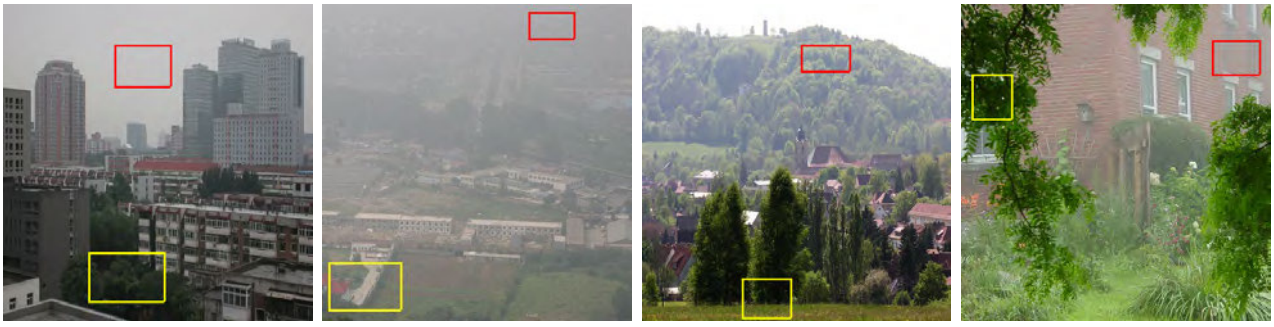


FIGURE 8. Demo of the degradations in the same image is not similar. We can see that the haze in yellow rectangles is thiner than areas in red ones.

RESIDE dataset simulates hazy images using NYU indoor dataset [33]. The indoor part of RESIDE dataset contains 500 hazy images for test. We then evaluate the performance of our proposed network on the SOTS dataset from

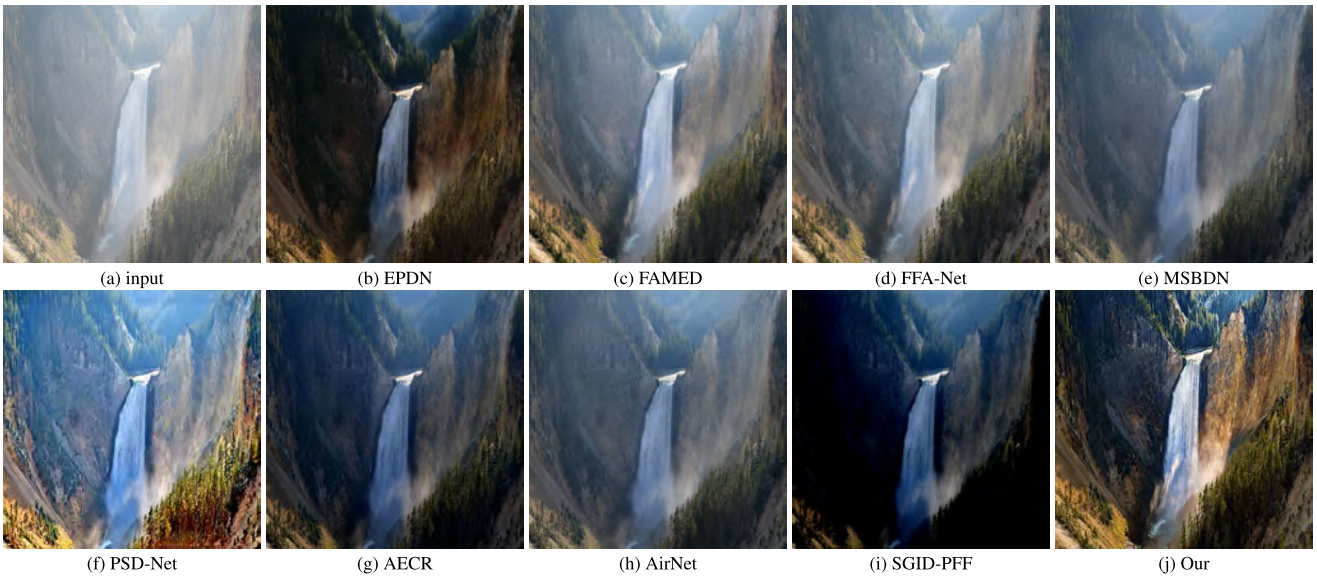


FIGURE 9. Visual results of some recently dehazing methods and proposed method.

RESIDE [28]. The comparison results on SOTS are shown in Table 1. From the experimental comparisons, it has been demonstrated that the proposed method outperforms the current state-ofthe-art methods [21], [29], and achieves superior performance with great improvements. It should be pointed out that the FFA-Net achieves the best scores for RESIDE dataset. However, its performance on real hazy images is poor. We term the GridDehazeNet [29] as GDN.

B. QUALITATIVE COMPARISON

To further evaluate the proposed method, we use real images to compare with different state-of-the-art methods. Fig. 6 shows the qualitative comparison of results with the seven state-of-the-art dehazing algorithms [2], [9], [11], [17], [18], [20], [21], [31], [32] on challenging real-world images. As shown in Fig. 6(b), most of the haze is removed by DCP, and the details of the scenes and objects are well restored. However, the results significantly suffer from over-enhancement (for instance, the building regions of the first and second images are much darker than it should be.

The results of DehazeNet, AOD-Net, FFA-Net, MSBDN, Dehamer, GridDehazeNet, and DCPDN do not have the over-estimation problem and maintain the original colors of the objects as shown in Fig. 6. But these methods have some remaining haze in the dehazed results. The method of AirNet and GFN tend to non-uniformly estimate haze concentration and results in inhomogeneous dehazed images in Fig. 6(k). The PhysicsGan, and EPDN generates relatively clear results, but the images show some color distortions. In contrast, the dehazed results by our method are clear and the details of the scenes are enhanced moderately as shown in Fig. 6(n).

We note some works employ hazy images to train dehazing network [6], [31]. PSD [6] employ hazy images to fine-tune dehazing network. AirNet [6] designs a encoder-decoder network, which can handle unknown corruption images. As shown in Fig. 7, we can see that the dehazed of AirNet and PSD show a haze appearance. We also note some color distortion in the deazed result of PSD in the second row in Fig. 7. The dehazed of AirNet looks darker than the dehazed results of proposed method and PSD. In contrast, the proposed method can restore the images details and recover a reasonable global appearance. The AirNet assumes that the degradations in the same image should be similar, which is not true for image dehazing. We show an example in Fig. 8, which shows that the degradations in the same image is not similar. The PSD employs several well-grounded physical priors to fine-tune the dehazing model. However, the physical priors arenot true for all hazy images. The proposed method employ haze-aware model to fuse the dehazed result, which helps the model restore high quality dehazed result.

We further compare the proposed method with some recently End-to-End dehazing methods [5], [6], [21], [30], [31], [34], [35], [36]. We show an example in Fig. 9. The dehazed results of FAMED, FFA-Net, MSBDN, AECR, and AirNet tend to retain haze. EPDN can remove haze. However, the dehazed result of EPDN tend to lose image details. SGID-PFF can remove haze. However, some areas of dehazed are completely dark. PSD can enhance the hazy image. However, the dehazed result of PSD tends to retain haze and show color distortion. In contrast, the proposed method can restore the images details and recover a reasonable global appearance and colorful dehazed result.

To show the effectiveness of the proposed method, we compare it with other dehazing methods. First, we show the dehazing performance of dehazing methods on real hazy images. Second, we show the densities of dehazed results obtained by different dehazing methods. As shown in Table 2, we can see that the proposed method achieve the second best

TABLE 4. Quantitative comparison results on the indoor hazy images from RESIDE with different modules and losses.

	Base	BaseNet	BaseAFM	BaseFMB	L_1	Full
PSNR	27.05	32.89	33.26	34.03	34.60	34.80
SSIM	0.92	0.97	0.98	0.98	0.98	0.99

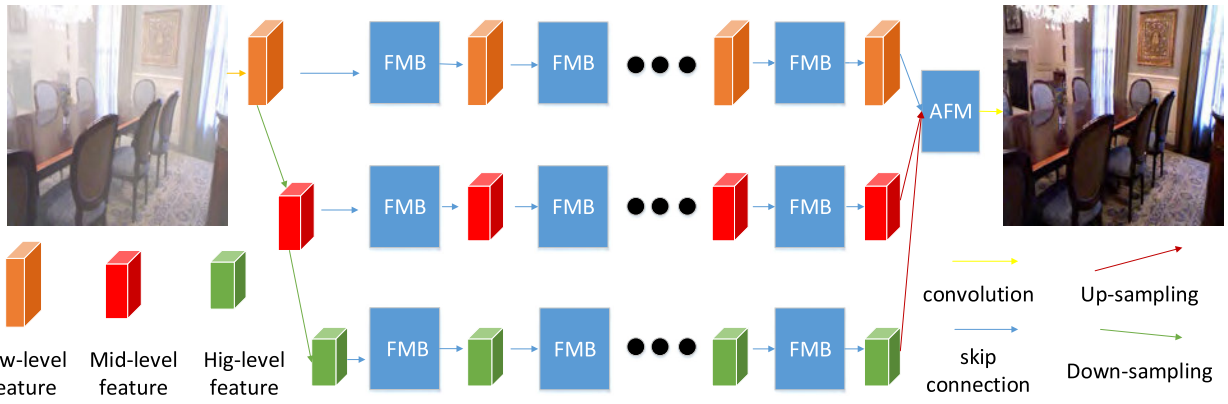


FIGURE 10. The architecture of the BaseFMB. As shown, BaseFMB contains three scales, and FMBs are used to extract features at each scale, AFM represents the adaptive fusion module.

dehazing performance with metric DHQI [37]. The proposed method is a data-driven dehazing method, which may do not perform well for real-world images. However, the proposed method is better than other data-driven dehazing methods. To Further show the performance of the proposed method, we show the density of the results obtained by dehazing methods. As shown in Table 3, we can see that the proposed method can remove haze better than other dehazing methods.

C. ABLATION STUDY

To better show the effectiveness of the proposed modules, we design an ablation study that includes a coarse multi-scale network, fine multi-scale blocks, and adaptive fusion modules. We construct the series variants with different proposed modules: 1) To show the effectiveness of the coarse multi-scale network, we design a single scale model termed as Base; 2) We add coarse multi-scale ability by removing AFM and replacing FMBs with traditional dense blocks, which is termed as BaseNet. We show the architecture of the BaseFMB in Fig. 10; 3) We add AFM to the BaseNet, and we term it as BaseAFM; 4) We replace traditional dense blocks with FMBs, and we term it as BaseFMB; 5) The architecture proposed in Section III, which is termed as Full. All models are trained in the same way and tested on the indoor part from RESIDE. As shown in Table 4, each proposed module shows its contribution to image dehazing.

To show the influence of loss function, we add an experiment, which only use L_1 norm to train the proposed model. As shown in Table 4, we can see that model trained with L_1 norm obtains lower quality dehazing results.

To show the efficiency of the proposed model, we show the run time from the variants of the proposed model. As shown in Table 5, we can see that the proposed model run faster than other models, such as Base, BaseNet, and BaseAFM. The models are tested on a computer equipped with a Nvidia Geforce 1060.

TABLE 5. Run time comparison results for the variants of the proposed model on the indoor hazy images.

method	Base	BaseNet	BaseAFM	BaseFMB	Full
time (ms)	0.0303	0.0309	0.0310	0.0288	0.0289

TABLE 6. Run time comparison results for the proposed method and other state-of-the-art dehazing methods on the indoor hazy images.

method	GDN	MSDBN	DMSDN
time (ms)	268.82	199.14	111.80

D. RUN TIME

We note that the dehazing performance has been greatly improved. However, the dehazing speed is slow. In this subsection, we compare the propose model with some dehazing methods, which achieve high dehazing performance. We test the dehazing speed on a server platform, which is equipped with eight TITAN V GPUs. The CPU of the platform is Intel(R) Xeon(R) Gold 6132 CPU @ 2.60GHz and the memory is 512 GB. We resize the hazy images to a fix size 512 × 512. We show the dehazing speed of state-of-the-art methods in Table 6. As we can see that the proposed dehazing is almost two times faster than MSDBN.

E. LIMITATION

Although the proposed model is effective for most hazy images. However, the proposed method maybe failed for

FIGURE 11. Visual results of the transmission maps and the corresponding dehazed result of guided with (b).

some dense hazy images. We address the this problem by using DCP loss. We show an example in Fig. 11, which is from the prior work [38]. As shown, we can see that the dcp loss may result in artifacts around depth jump areas. In order to further improve the dehazing quality, we design a novel method to improve the accuracy of transmission map. We show the different between the DCP and the proposed method in Fig. 11. Although the proposed model is trained with the transmission maps estimated by DCP, The proposed model also trained with synthetic dataset, which improve the accuracy of the estimated transmission maps. As shown in Fig. 11, we can see that the the transmission maps estimated by DCP contains more details. In contrast, the visual result of the proposed method is much smoother. We can use the new transmission map predicting network and real hazy image to boost the dehazing performance on real hazy images.

As proved by [39], [40], DNN-based methods often learn low-frequency functions, while ignore the high-frequency information. The neural augmentation framework [24] is proposed to address such a problem. In the feature, we also adopt the neural augmentation framework to improve the dehazing quality of the proposed method.

V. CONCLUSION

In this paper, we design a dual multi-scale dehazing network for single-image dehazing. The model contains a coarse multi-scale network and fine multi-scale blocks. The coarse multi-scale network which capture is designed to capture large variations of object sizes, while fine multi-scale blocks are designed to capture small variations of object sizes. The coarse multi-scale network contains three scales, which extract pyramidal features from the input image. To further explore the multi-scale information, we develop a fine multi-scale block, which extracts multi-scale information using dilation convolution with different dilation rates and channel-wise attention. The adaptive fusion module is designed to boost information flowing. Extensive experiments are conducted on public synthetic indoor images and natural hazy images to show the effectiveness of the proposed method.

REFERENCES

[1] E. J. McCartney, *Optics of the Atmosphere: Scattering by Molecules and Particles*. New York, NY, USA: Wiley, 1976, p. 421.

[2] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341–2353, Dec. 2011.

[3] 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.

[4] A. Singh, A. Bhave, and D. Prasad, "Single image dehazing for a variety of haze scenarios using back projected pyramid network," in *Proc. Eur. Conf. Comput. Vis.*, vol. 8 2020, pp. 166–181.

[5] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Xie, "FFA-Net: Feature fusion attention network for single image dehazing," in *Proc. AAAI Conf. Artif. Intell.*, Feb. 2020, pp. 11908–11915.

[6] Z. Chen, Y. Wang, Y. Yang, and D. Liu, "PSD: Principled synthetic-to-real dehazing guided by physical priors," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 7176–7185.

[7] Y. Liu, Z. Yan, J. Tan, and Y. Li, "Multi-purpose oriented single nighttime image haze removal based on unified variational retinex model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 4, pp. 1643–1657, Apr. 2023.

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

[9] J. Pan, J. Dong, Y. Liu, J. Zhang, J. Ren, J. Tang, Y.-W. Tai, and M.-H. Yang, "Physics-based generative adversarial models for image restoration and beyond," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 7, pp. 2449–2462, Jul. 2021.

[10] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 617–624.

[11] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "DehazeNet: An end-to-end system for single image haze removal," *IEEE Trans. Image Process.*, vol. 25, no. 11, pp. 5187–5198, Nov. 2016.

[12] Y. Li, S. You, M. S. Brown, and R. T. Tan, "Haze visibility enhancement: A survey and quantitative benchmarking," *Comput. Vis. Image Understand.*, vol. 165, pp. 1–16, Dec. 2017.

[13] R. Fattal, "Single image dehazing," *ACM Trans. Graph.*, vol. 27, no. 3, p. 72, Aug. 2008.

[14] K. Tang, J. Yang, and J. Wang, "Investigating haze-relevant features in a learning framework for image dehazing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2995–3002.

[15] Q. Zhu, J. Mai, and L. Shao, "A fast single image haze removal algorithm using color attenuation prior," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3522–3533, Nov. 2015.

[16] 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.*, 2016, pp. 154–169.

[17] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "An all-in-one network for dehazing and beyond," in *Proc. IEEE Int. Conf. Comput. Vis.*, Aug. 2017, pp. 1–12.

[18] H. Zhang and V. M. Patel, "Densely connected pyramid dehazing network," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3194–3203.

[19] X. Yang, Z. Xu, and J. Luo, "Towards perceptual image dehazing by physics-based disentanglement and adversarial training," in *Proc. AAAI Conf. Artif. Intell.*, 2018, pp. 1–8.

[20] W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W. Liu, and M.-H. Yang, "Gated fusion network for single image dehazing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3253–3261.

[21] H. Dong, J. Pan, L. Xiang, Z. Hu, X. Zhang, F. Wang, and M.-H. Yang, "Multi-scale boosted dehazing network with dense feature fusion," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 2154–2164.

[22] Z. Zheng, W. Ren, X. Cao, X. Hu, T. Wang, F. Song, and X. Jia, "Ultra-high-definition image dehazing via multi-guided bilateral learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 16180–16189.

[23] Y. Shao, L. Li, W. Ren, C. Gao, and N. Sang, "Domain adaptation for image dehazing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 2805–2814.

[24] Z. Li, C. Zheng, H. Shu, and S. Wu, "Dual-scale single image dehazing via neural augmentation," *IEEE Trans. Image Process.*, vol. 31, pp. 6213–6223, 2022.

[25] Z. Li, C. Zheng, H. Shu, and S. Wu, "Single image dehazing via model-based deep-learning," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2022, pp. 141–145, doi: 10.1109/ICIP46576.2022.9897479.
[26] C. Zheng, Z. Li, Y. Yang, and S. Wu, "Single image brightening via multi-scale exposure fusion with hybrid learning," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 4, pp. 1425–1435, Apr. 2021, doi: 10.1109/TCSVT.2020.3009235.
[27] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," in *Proc. Int. Conf. Learn. Represent.*, 2016, pp. 1–13.
[28] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang, "Benchmarking single-image dehazing and beyond," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 492–505, Jan. 2019.
[29] X. Liu, Y. Ma, Z. Shi, and J. Chen, "GridDehazeNet: Attention-based multi-scale network for image dehazing," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 7313–7322.
[30] Y. Qu, Y. Chen, J. Huang, and Y. Xie, "Enhanced Pix2pix dehazing network," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 8152–8160.
[31] B. Li, X. Liu, P. Hu, Z. Wu, J. Lv, and X. Peng, "All-in-one image restoration for unknown corruption," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, New Orleans, LA, Jun. 2022, pp. 17431–17441.
[32] C. Guo, Q. Yan, S. Anwar, R. Cong, W. Ren, and C. Li, "Image dehazing transformer with transmission-aware 3D position embedding," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 5802–5810.
[33] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from RGBD images," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 746–760.
[34] J. Zhang and D. Tao, "FAMED-Net: A fast and accurate multi-scale end-to-end dehazing network," *IEEE Trans. Image Process.*, vol. 29, pp. 72–84, 2020.
[35] H. Bai, J. Pan, X. Xiang, and J. Tang, "Self-guided image dehazing using progressive feature fusion," *IEEE Trans. Image Process.*, vol. 31, pp. 1217–1229, 2022.
[36] H. Wu, Y. Qu, S. Lin, J. Zhou, R. Qiao, Z. Zhang, Y. Xie, and L. Ma, "Contrastive learning for compact single image dehazing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2021, pp. 10551–10560.
[37] X. Min, G. Zhai, K. Gu, X. Yang, and X. Guan, "Objective quality evaluation of dehazed images," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 8, pp. 2879–2892, Aug. 2019, doi: 10.1109/TITS.2018.2868771.
[38] Z. Li, H. Shu, and C. Zheng, "Multi-scale single image dehazing using Laplacian and Gaussian pyramids," *IEEE Trans. Image Process.*, vol. 30, pp. 9270–9279, 2021, doi: 10.1109/TIP.2021.3123551.
[39] R. Basri, D. W. Jacobs, Y. Kasten, and S. Kritchman, "The convergence rate of neural networks for learned functions of different frequencies," in *Proc. Annu. Conf. Neural Inf. Process. Syst.*, Vancouver, BC, Canada, Dec. 2019, pp. 4763–4772.
[40] N. Rahaman, A. Baratin, D. Arpit, F. Draxler, M. Lin, F. A. Hamprecht, Y. Bengio, and A. C. Courville, "On the spectral bias of neural networks," in *Proc. 36th Int. Conf. Mach. Learn. (ICML)*, Long Beach, CA, USA, Jun. 2019, pp. 5301–5310. [Online]. Available: http://proceedings.mlr.press/v97/rahaman19a.html

SHENGDONG ZHANG received the B.S. degree in computer science and technology from Shanxi University, Shanxi, China, in 2009, the M.S. degree in mechanics from Peking University, Beijing, China, in 2012, and the Ph.D. degree from Wuhan University, Wuhan, China, in 2020. His research interests include deep learning, computer vision, and image processing.

XIAOQIN ZHANG received the B.Sc. degree in electronic information science and technology from Central South University, China, in 2005, and the Ph.D. degree in pattern recognition and intelligent system from the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, China, in 2010. He is currently a Professor with Wenzhou University, China. His research interests include pattern recognition, computer vision, and machine learning. He has published more than 100 papers in international and national journals, and international conferences, including IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, *IJCV*, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, ICCV, CVPR, NIPS, IJCAI, AAAI, ACM MM, and among others.

LINLIN SHEN (Senior Member, IEEE) received the B.Sc. and M.Eng. degrees from Shanghai Jiaotong University, Shanghai, China, and the Ph.D. degree from the University of Nottingham, Nottingham, U.K. He is currently a Professor in computer science and software engineering with Shenzhen University, Shenzhen, China. He is also an Honorary Professor with the School of Computer Science, University of Nottingham, U.K. He serves as the Director of the Computer Vision Institute and the China–U.K. Joint Research Laboratory for Visual Information Processing. He was a Research Fellow with the University of Nottingham, working on MRI brain image processing. His research interests include deep learning, facial recognition, analysis/synthesis, and medical image processing. He is listed as the Most Cited Chinese Researcher by Elsevier. He received the Most Cited Paper Award from the journal of *Image and Vision Computing*. His cell classification algorithms were the winners of the International Contest on Pattern Recognition Techniques for Indirect Immunofluorescence Images held by ICIP 2013 and ICPR 2016.