IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 13, 2020, accepted January 27, 2020, date of publication February 3, 2020, date of current version February 10, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2971092

MMP-Net: A Multi-Scale Feature Multiple Parallel Fusion Network for Single Image Haze Removal

JIAJIA YAN[1], CHAOFENG LI[1], (Senior Member, IEEE), YUHUI ZHENG[2], SHOUKUN XU[3], AND XIAOYONG YAN[4,5]

[1]Institute of Logistics Science and Engineering, Shanghai Maritime University, Shanghai 201306, China
[2]College of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, China
[3]School of Information Science and Engineering, Changzhou University, Changzhou 213164, China
[4]School of Modern Posts and Institute of Modern Posts, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
[5]Key Laboratory of Applied Mathematics (Putian University), Fujian Province University, Putian 351100, China

Corresponding author: Chaofeng Li (wxlichaofeng@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61771223, and in part by the Key Laboratory of Applied Mathematics of Fujian Province University (Putian University) under Grant SX201905.

ABSTRACT Reducing the impact of hazy images on subsequent visual information processing is a challenging problem. In this paper, combining with atmospheric scattering model, we propose an end-to-end multi-scale feature multiple parallel fusion network called MMP-Net for single image haze removal. The MMP-Net includes three components: multi-scale CNN module, residual learning module and deep parallel fusion module. 1) In multi-scale CNN module, a multi-scale convolutional neural network (CNNs) is adopted to extract different scales features from whole to local, and these features are fused multiple times in parallel. 2) In residual learning module, residual blocks are introduced to deeply learn detailed features, which can recover more image details. 3) In deep parallel fusion module, those features from residual learning module are deeply merged with the fused features from CNNs, and finally used to recover a clean haze-free image via the atmospheric scattering model. The experimental results show that on the average of three datasets (SOTS, HSTS, and D-Hazy), proposed MMP-Net improves PSNR from 20.91db to 22.21db and SSIM from 0.8720 to 0.9023 over the best state-of-the-art DehazeNet method. What's more, MMP-Net gains the best subjective visual quality on real-world hazy images.

INDEX TERMS Image dehazing, convolutional neural network, residual learning, parallel fusion.

I. INTRODUCTION

In hazy weather, due to the fine particles suspended in the atmosphere, the outdoor image captured by the machine is scattered, which causes a decline in image quality and a dim color of the image. This not only has a negative impact on human perception, but also constitutes an obstacle for many computer vision tasks, such as video surveillance [1], target recognition [2], [3], image classification [4], [5] and so on. Therefore, in order to improve the image quality and performance of computer vision systems, image haze removal has become an important research topic in current computer vision [6], [7].

At present, image dehazing methods can be summarized into three categories: one based on prior knowledge, one based on learning and the other based on a combination of prior knowledge and learning. For the first kind, researchers need to find the common features between hazy images and clean images. There are some image dehazing methods based on prior knowledge. Tan [8] found that the hazy image contrast is often weaker than the clean image, and propose a local contrast pair maximization algorithm. However, increasing the contrast of the image for dehazing may cause image color distortion and loss of image realism. He *et al.* [9] assumed that haze-free images often have low-intensity values in at least one channel. Therefore, a dark channel prior dehazing algorithm is proposed to remove haze by an atmospheric scattering model. However, color distortion is likely to occur in the sky area and other areas that do not satisfy the dark channel prior. Zhu *et al.* [10] discovered the linear relationship between scene depth, brightness and saturation of images, and proposed a color attenuation prior method.

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

Berman and Avidan [11] found that the colors of a clean image form a tight cluster in the RGB space. Then, a non-local prior image dehazing method is proposed. Kratz and Nishino [12] used the factorial MRF model to estimate albedo and depth, and factorized a single hazy image via a canonical expectation-maximization algorithm with alternating minimization. In addition, the prior knowledge-based method can also be used for images captured at night and under non-uniform illumination. Zhang [13] found that ambient illumination is the main source of local maximum intensity for each color channel in nighttime hazy images. Subsequently, a maximum reflectance prior is proposed to estimate the ambient illumination and transmission map. Although methods based on prior knowledge are simple and effective in many scenarios, they still have some limitations in specific scenarios. What's more, it is quite difficult to extract image features and prior knowledge artificially.

Different from the prior knowledge-based method, the learning-based image dehazing method no longer relies on manually extracting image features, but automatically obtains image features by learning model. In recent years, many researchers have studied learning-based image dehazing methods. Cai *et al.* [14] proposed a DehazeNet with the new BReLU activation function to dehaze by predicted transmittance map. Ren *et al.* [15] used a multi-scale CNN to predict the transmission map for image dehazing. Li *et al.* [16] proposed a deep all-in-one model to generate clean images directly by convolutional neural networks. Zhang and Patel [17] proposed an end-to-end densely connected pyramid network (DCPDN) to jointly learn the transmission map, atmospheric light. In addition, a joint discriminator based on a generative adversarial network framework is used to determine whether the estimated transmission map is real or fake. In [18] a Cycle-Dehaze network was proposed for single image dehazing, which does not rely on the estimation of atmospheric scattering model, but instead produces visually better haze-free images by improving the quality of texture information recovery. Zhang *et al.* [19] proposed a fully point CNN (FPCNet) method. This network shuffles the original images and inputs them to the network, which can model the statistical regularities effectively. Finally, the haze is removed by point-wise convolutions in all convolutional layers. Du and Li [20] proposed to redefine dehazing as a problem of learning structural residue. A deep residual learning (DRL) network that directly estimates the non-linear mapping of the input image to the output image is designed for dehazing. Mei *et al.* [21] proposed a progressive feature fusion network to directly learn a non-linear function from a hazy image to a clean image. Although it can recover hazy images with 4K resolution, dehazed images exhibit noise such as grid shapes. Zhang and Tao [22] proposed a network consisting of encoders at three scales and a fusion module. This multi-scale end-to-end dehazing network can quickly and accurately learn haze-free images, which called FAMED-Net. Dudhane and Murala [23] proposed RNet and YNet to extract haze-related features through RGB and YCbCr color spaces and generate two transmission maps (TrMaps). The TrMap Converged Network (FNet) is designed to integrate two TrMaPs. This network structure is called RYF-Net. Currently, learning-based methods can also be applied to night images. Kuanar *et al.* [24] proposed a deep learning based DeGlow-DeHaze iterative architecture. The DeGlow model eliminates the glow generated in the night scene, and the DeHaze network eliminates the haze. Although improvements have been gained by the learning-based methods, their dehazed performances are still hindered by the estimation of the transmittance map and atmospheric light and deep neural network structure. The current dehazing methods still leave much space to improve.

In recent years, some dehazing methods based on a combination of prior knowledge and learning have also been born. Ren *et al.* [25] presented an end-to-end gating fusion network consisting of an encoder and a decoder for dehazing, where image preprocessing is applied via white balance (WB), contrast enhancement (CE) and gamma correction (GC). Wang *et al.* [26] found that haze mainly affects the luminance channel in YCrCb color space and proposed a novel simple but powerful atmospheric illumination prior (AIP). Based on this prior knowledge, a multi-scale convolutional neural network that automatically recognizes hazy areas, and restoration of deficient texture information is designed for image dehazing. These networks cleverly combine prior knowledge and succeed in dehazing a single image. However, there is still extensive optimization that needs further investigate.

In this paper, we propose a MMP-Net for single image haze removal. Three different scales features are extracted by a multi-scale CNN model and fused multiple times in parallel. Further, the residual learning blocks are introduced to get more detailed image features. These features are deeply fused with those from CNNs in parallel, and finally utilized to produce the dehazed image via atmospheric scattering model.

Here we firstly give the visual quality subjective comparison between MMP-Net and several state-of-the-art methods on a natural hazy image shown as Fig.1. It can be seen these models can augment image details and enhance visual quality. However, color distortion appears in Fig.1 (b), and dehazed image shows not natural. In Fig.1 (c) and Fig.1 (d), the texture is not clear, and dehazed images shows a little blurred. Compared with Fig.1 (e), Fig.1 (f) shows clearer texture and more realistic colors. In general, Fig.1 (f) shows the fewest color distortions and the best visual quality.

Further we also give the objective assessment results of Fig.1 by using blind image quality assessment metric NIQE [27], shown in subtitle of Fig.1. A lower NIQE indicates better image quality. It can be seen our MMP-Net gets the lowest NIQE value, suggesting the best dehazed performance.

Our main contributions can be summarized as follows.

(1) We propose an end-to-end multi-scale feature multiple parallel fusion network MMP-Net, which consists of multi-scale CNN module, residual learning module and deep parallel fusion module.

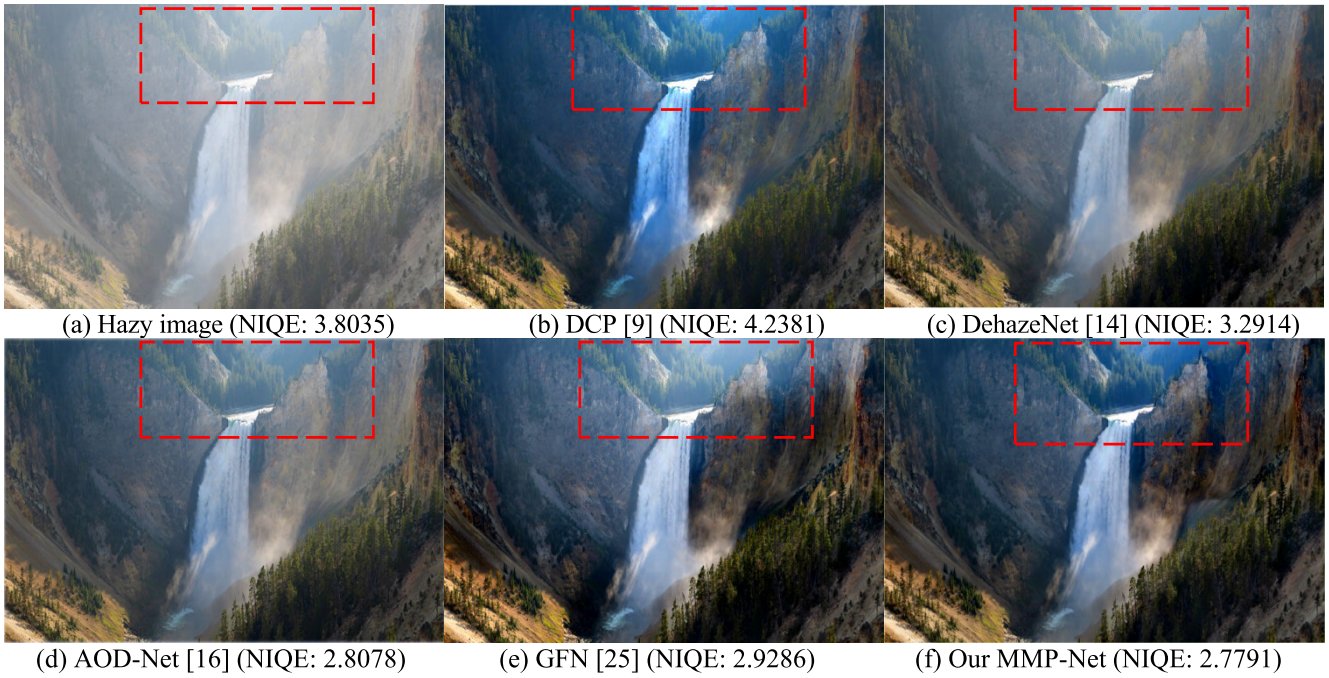


FIGURE 1. Visual quality comparison between MMP-Net and several state-of-the-art methods on a natural hazy image. Please amplify figures to view the detail differences in bounded regions. Note that the score of NIQE are range from 0 (the best) to 100 (the worst).

(2) In MMP-Net, the residual blocks are introduced to deeply learn detailed features, and different scales features are deeply merged multiple times in parallel.

(3) Experimental results show proposed MMP-Net outperformed current reported top image dehazing methods both on synthetic SOTS, HSTS and D-Hazy hazy image datasets and real-world hazy images.

II. ATMOSPHERIC SCATTERING MODEL AND TRANSFORMED EXPRESSION

McCartney [28] proposed an atmospheric scattering model for expressing haze image formation as follows:

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

where $I(x)$ is the observed hazy image; $J(x)$ is the scene radiance (i.e., the ideal "clean image") to be recovered; A is the global atmospheric light; $t(x)$ is the haze transmission matrix; and x denotes pixel location. $t(x)$ is defined as:

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

where β denotes the medium attenuation coefficient, and $d(x)$ is the scene depth.

The dehazing process is to estimate the transmittance map $t(x)$ and the atmospheric light A from the hazy image $I(x)$. Then, the Eq. (1) can be rewritten into an expression about a clean image:

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

From Eq. (3) it can be seen the transmittance map $t(x)$ and the atmospheric light A have a great influence on dehazing effect. For a single image haze removal, the haze image $I(x)$ is known. However, the A and $t(x)$ are unknown. For simplifying the two unknowns, according to [16], a single variable $K(x)$ will be used to get a new mathematical expression as follows.

$$J(x) = K(x)I(x) - K(x) + b \tag{4}$$

where b is the constant bias with the default value 1. Then $K(x)$ can be expressed as following Eq. (5).

$$K(x) = \frac{\frac{1}{t(x)}(I(x) - A) + (A - b)}{I(x) - 1} \tag{5}$$

After $K(x)$ is predicted by a deep learning model, and then a clean image $J(x)$ can be gained.

III. PROPOSED MMP-Net FOR SINGLE IMAGE DEHAZING

In this section, we introduce our proposed MMP-Net, which includes three parts: multi-scale CNN module, residual learning module and deep parallel fusion module. The multi-scale CNN module consists of 6 convolutional layers, each followed by a nonlinear ReLU activation. The residual learning module consists of 18 residual blocks, each of which contains a PReLu activation function. The deep parallel fusion module is used to further fuse different scales features from CNN and residual learning module. The schematic diagram of a single image dehazing via MMP-Net is shown as Fig.2.

A. MULTI-SCALE CNN MODULE

Firstly, the multi-scale CNN module (the 1st part in Fig.2) uses three convolution kernels of different scales to extract

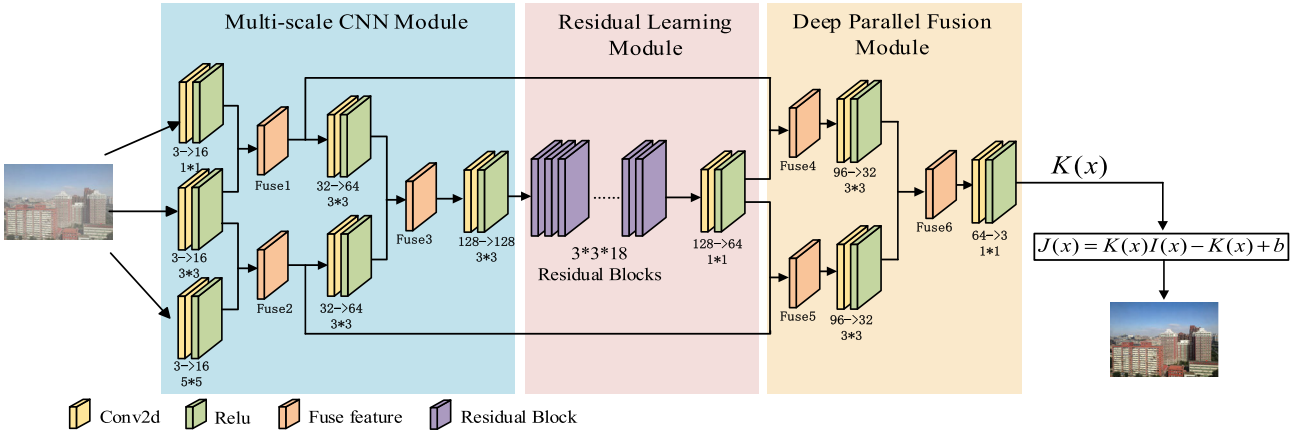


FIGURE 2. Schematic diagram of a single image dehazing via MMP-Net.

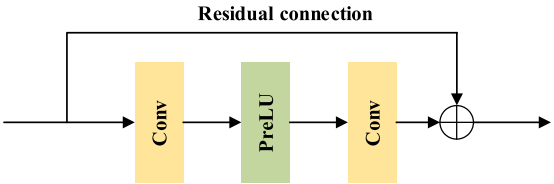


FIGURE 3. Residual block structure.

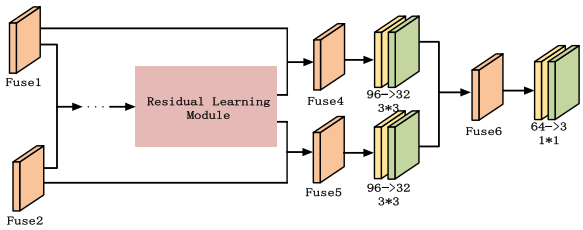


FIGURE 4. Architecture of deep feature parallel fusion module.

features from the hazy image. Here, the features obtained by the convolutional kernels of 1×1 and 3×3 are merged in parallel to obtain Fuse1. The 3×3 and 5×5 convolutional kernels get Fuse2 by the same operation. Fuse1 and Fuse2 are then convoluted and merged in parallel to obtain Fuse3. The $1\times$ convolutional kernel is used to preserves as much detail as possible of the original image, especially edge information. Multiple parallel fusions are used to compensate lost features by 3×3 and 5×5 convolution kernels. Multi-scale CNN module facilitates the transfer of detailed features obtained via small-scale convolution into subsequent images to generate rich features and edge information. The feature extraction by convolution operation is expressed as follows:

$$G_i(Y) = W_i * G_{i-1}(Y) \tag{6}$$

where $G_i(Y)$ denotes the feature map of the $i-$ th layer of the output; W_i is the convolution kernel of the $i-$ th layer; $G_{i-1}(Y)$ is the feature map of the $(i-1)-$ th layer; and $*$ represents a convolution operation.

B. RESIDUAL LEARNING MODULE

In MMP-Net, we introduce residual learning [29] (the 2nd part in Fig.2) into a deep network structure for extracting more detailed image features. Mei *et al.* [21] show that 18 residual blocks can keep the balance between learning performance and computing resources, so we also use 18 residual blocks in the MMP-Net. The skip connections are adopted between the residual blocks, which can reduce the problem of gradient disappearance caused by deep neural networks.

Fig.3 shows the structure of the residual block, which consists of 2 convolutional layers and a non-linear activation function PreLU. The authors of [30] suggest using PreLU can not only alleviate the vanishing gradient problem, but also speed up the convergence of the network, so it is chosen as the activation function of MMP-Net. The PreLU is expressed as:

$$f_{\mathrm{PReLU}} = \max(x_i, 0) + \alpha_i \min(0, x_i) \tag{7}$$

where x_i is the input of the nonlinear activation f on the $i-$ th layer, and α_i is a coefficient controlling the slope of the negative part.

C. DEEP PARALLEL FUSION MODULE

Fig.4 shows the architecture of deep feature parallel fusion module. In this module, the features from residual learning module are merged with Fuse1 and Fuse2 separately in parallel to get Fuse4 and Fuse5. Fuse4 and Fuse5 are convolved separately and then merged to get Fuse6 in parallel. Fuse6 is convolved to get the final feature map $K(x)$. At last, dehazed image $J(x)$ is got by taking $K(x)$ into Eq. (4).

IV. EXPERIMENTAL RESULTS AND ANYLYSIS

Dataset: NYU Depth Dataset V2 [31] is used as a training set, which includes 27,256 hazy images of different concentrations and 1449 clean images with size of 640×480. Specially, given a clean image J, random atmospheric light $A \in [0.7, 1.0]$ for each channel, and the corresponding

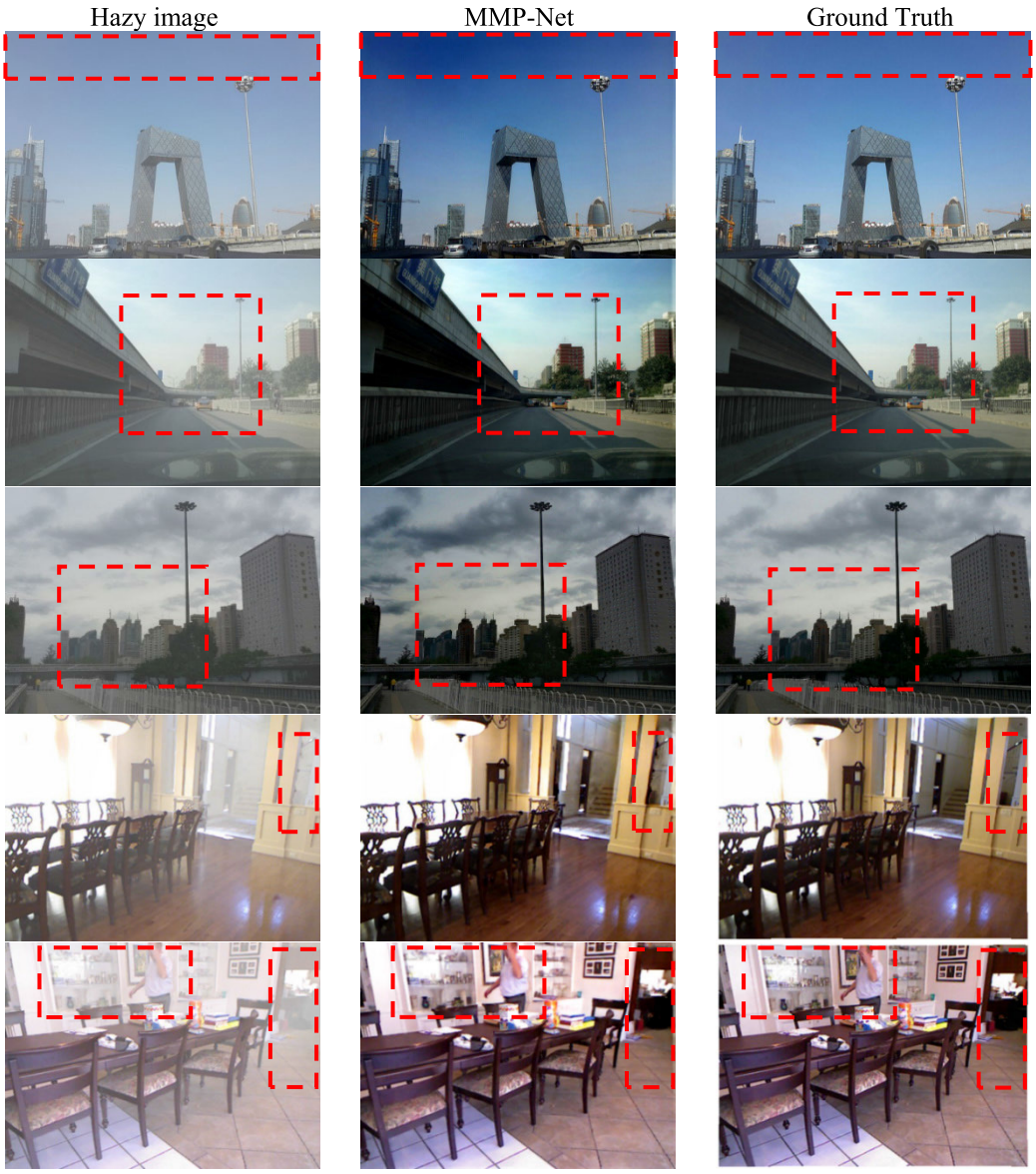


FIGURE 5. Dehazed results on the SOTS dataset.

ground-truth depth map d, function $t(x) = e^{-\beta d(x)}$ is applied to synthesize transmission map first, then a hazy image is generated by using the physical model in Eq.(1) with randomly selected scattering coefficient $\beta \in [0.6, 1.8]$.

Three datasets are selected as test sets to evaluate the dehazing performance. The Synthetic Objective Testing Set (SOTS) of RESIDE [32] includes both indoor and outdoor scenes with 500 images of each. The Hybrid Subjective Testing Set (HSTS) of RESIDE consists of 10 outdoor synthetic hazy images. The D-hazy dataset [33] is a standard dataset for image dehazing, which consists of 1,449 synthetic hazy images and their respective clean images.

Experimental Setup: MMP-Net is implemented in the pytorch [34] framework and runs on a workstation with a 3.2 GHz CPU, 32G RAM, and Nvidia GeForce GTX 1080 Ti GPUs. During training, ADAM is used as the optimizer. The initial learning rate is 0.001, the batch size of images is 8 and then the model is trained for 120 epochs on GPU.

Several top single image dehazing methods are selected for comparing our proposed model, which includes DCP [9], CAP [10], DehazeNet [14], MSCNN [15], AOD-Net [16], GFN [25] and PFF-Net [21]. The DCP and CAP are based on prior knowledge, DehazeNet, MSCNN, AOD-Net, PFF-Net and RYF-Net are learning-based dehazing methods, and GFN [25] is based on a combination of prior knowledge and learning. Experiments are performed both on the synthetic SOTS, HSTS and D-Hazy hazy image dataset and the real-world hazy images dataset. Subsequently, the full-reference algorithm is used to evaluate the synthetic hazy image dataset, and the blind image quality assessment is used to evaluate the real-world hazy images dataset.

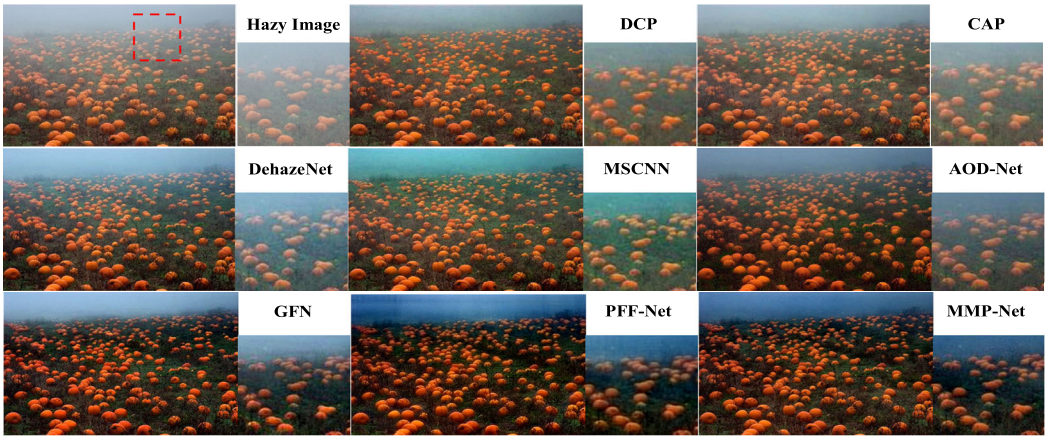


FIGURE 6. Subjective comparisons between MMP-Net and state-of-the-art methods on real-world hazy images.

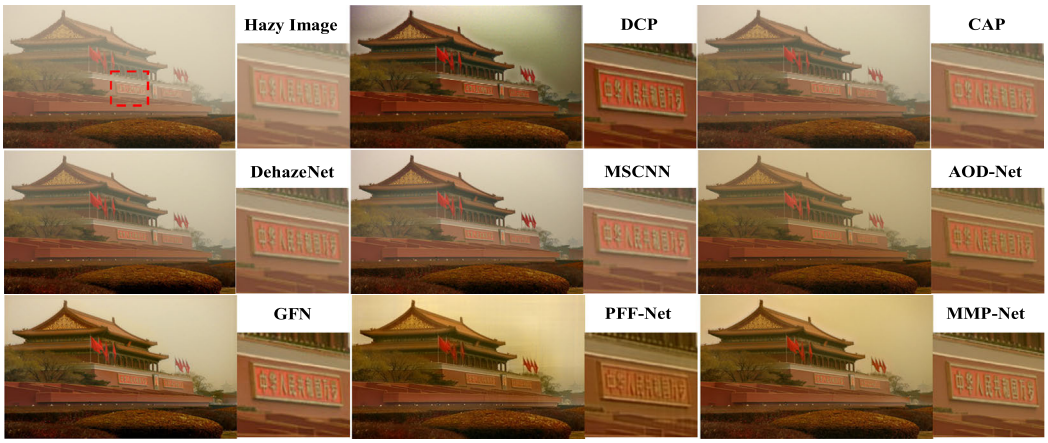


FIGURE 7. Subjective comparisons between MMP-Net and state-of-the-art methods on real-world hazy images.

A. RESULTS ON SYNTHETIC HAZY IMAGES

Fig.5 shows hazy images, the MMP-Net dehazed results and Ground Truth on the SOTS dataset. It can be seen MMP-Net dehazed images show very good subjective visual quality, although some deep and uneven color in the sky occurs in the dehazed images of first group, and some noise is distributed in the dehazed images of fourth group in the form of particles around the stairs.

We also give the objective assessment results listed in Table 1. The popular PSNR [35] and SSIM [36] are used to evaluate the dehazed performance. The higher the PSNR means that the image is less affected by noise. The SSIM reflects the similarity of image structure information. The closer the SSIM is to 1, the smaller the difference in distortion.

As shown in Table 1, compared with other models, MMP-Net gets the highest PSNR and SSIM on the SOTS dataset (both indoor dataset and outdoor dataset), which suggests the best detailed recovery and the least noise impact are obtained by the MMP-Net.

The HSTS and D-Hazy datasets are selected to verify the generalization ability of proposed model, and the results are given in Table 2 and 3. Since the authors of GFN and RYF-Net did not give their results on HSTS, their results of are not listed in Table 2.

TABLE 1. PSNR and SSIM on the SOTS of RESIDE.

Model	PSNR(indoor/outdoor)	SSIM(indoor/outdoor)
DCP[9]	17.88 (16.62/19.13)	0.8164 (0.8179/0.8148)
CAP[10]	19.05(19.05/-)	0.8364(0.8364/-)
DehazeNet[14]	21.80 (21.14/22.46)	0.8493 (0.8472/0.8514)
MSCNN[15]	17.57 (17.57/-)	0.8102 (0.8102/-)
AOD-Net[16]	19.68 (19.06/20.29)	0.8635 (0.8504/0.8765)
GFN[25]	21.93 (22.30/21.55)	0.8622 (0.8800/0.8444)
RYF-Net[23]	21.44 (21.44/-)	0.8716 (0.8716/-)
PFF-Net[21]	22.91 (24.70/21.12)	0.8687 (0.8951/0.8422)
MMP-Net	**25.30 (27.53/23.08)**	**0.9376 (0.9580/0.9161)**

TABLE 2. PSNR and SSIM on the HSTS of RESIDE.

Model	PSNR	SSIM
DCP[9]	14.84	0.7609
CAP[10]	21.53	0.8726
DehazeNet[14]	**24.48**	**0.9153**
MSCNN[15]	18.64	0.8168
AOD-Net[16]	20.55	0.8973
PFF-Net[21]	15.78	0.5778
Proposed MMP-Net	**23.02**	**0.9116**

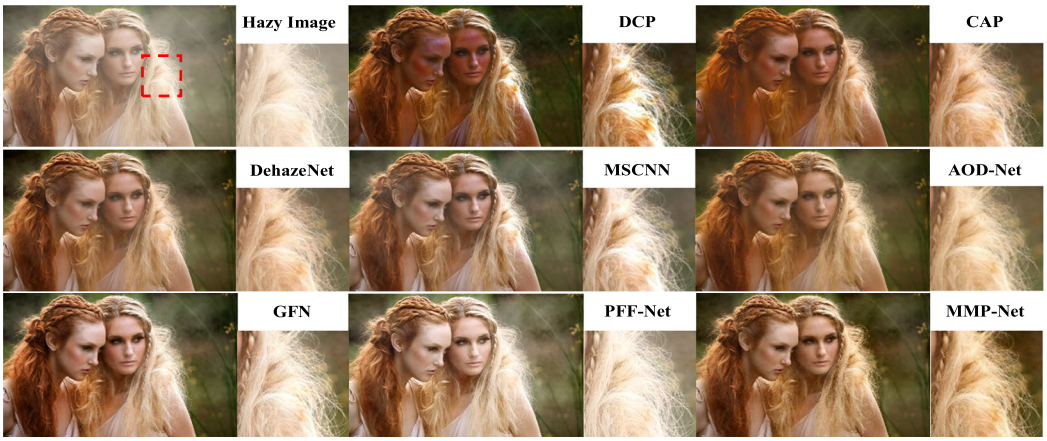


FIGURE 8. Subjective comparisons between MMP-Net and state-of-the-art methods on real-world hazy images.

FIGURE 9. Subjective comparisons between MMP-Net and state-of-the-art methods on real-world hazy images.

From Table 2 it can be seen on the HSTS dataset, the SSIM and PSNR of MMP-Net are only a little lower than of the DehazeNet. Owing to haze in all images of HSTS are thin, the residual structure of MMP-Net use skip connection, and maybe bypass some thin haze, which makes MMP-Net not shows the best on the HSTS.

Images in the D-Hazy dataset are with dense haze. From Table 3 it can be seen our proposed MMP-net get the best PSNR and SSIM on the D-hazy dataset, which suggests it is more effective in dense haze removal.

For comprehensive comparison, we give the average PSNR and SSIM of above algorithms on RESIDE, HSTS and D-Hazy datasets, shown as Table 4. It can be seen that proposed MMP-Net gets the highest PSNR and SSIM, and outperforms the state-of-art dehazing methods.

B. RESULTS ON REAL-WORLD HAZY IMAGES

In this subsection, we use 5 hazy images in natural scenes to further verify the reliability of the MMP-Net and dehazed results via MMP-Net and 7 state-of-the-art dehazing

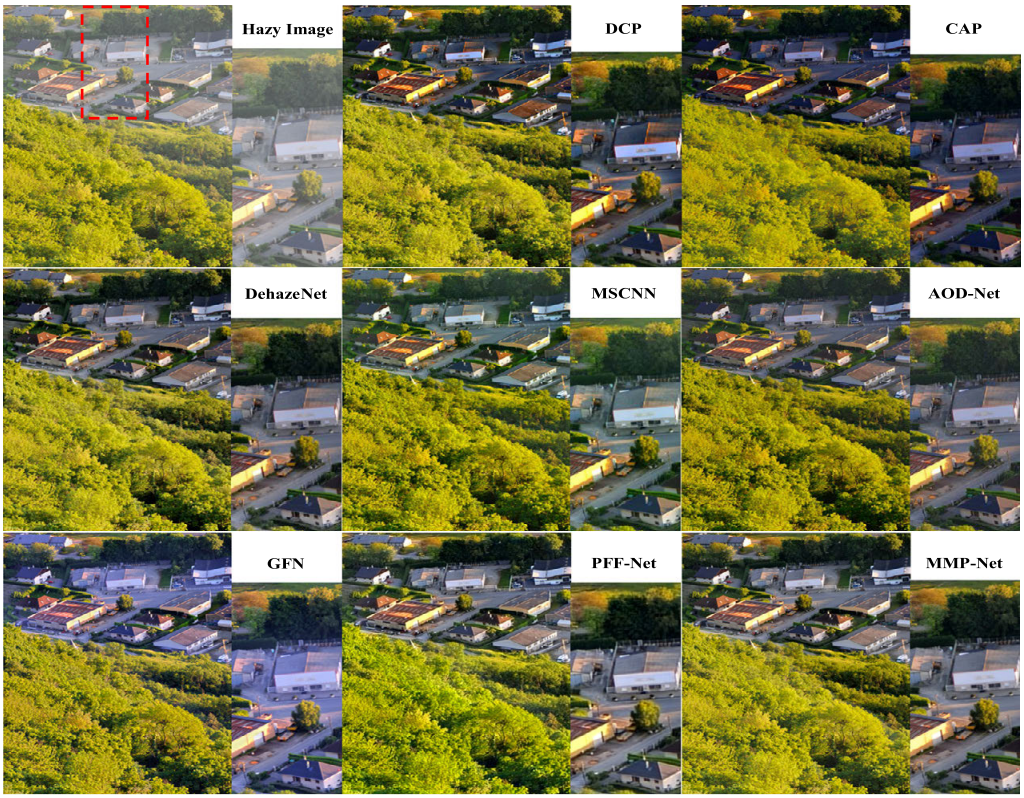
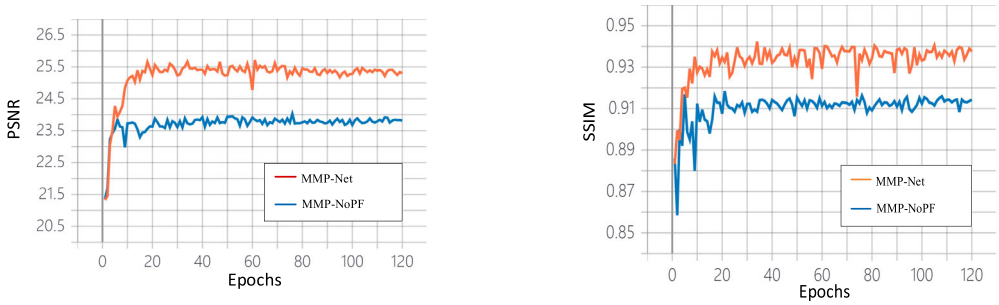


FIGURE 10. Subjective comparisons between MMP-Net and state-of-the-art methods on real-world hazy images.



(a) PSNR comparison curve (b) SSIM comparison curve

FIGURE 11. Comparison of the MMP-Net Variant on the SOTS dataset.

TABLE 3. PSNR and SSIM on the D-hazy.

Model	PSNR	SSIM
DCP[9]	15.49	0.8558
CAP[10]	15.17	0.8084
DehazeNet[14]	16.45	0.8513
MSCNN[15]	15.81	0.8443
AOD-Net[16]	12.44	0.7354
RYF-Net[23]	17.20	**0.8621**
PFF-Net[21]	16.64	0.7768
Proposed MMP-Net	**18.31**	**0.8576**

TABLE 4. Average PSNR and SSIM on three datasets.

Model	PSNR	SSIM
DCP[9]	16.07	0.8110
CAP[10]	18.58	0.8391
DehazeNet[14]	20.91	0.8720
MSCNN[15]	17.34	0.8238
AOD-Net[16]	17.6	0.8321
PFF-Net[21]	18.44	0.7411
Proposed MMP-Net	**22.21**	**0.9023**

algorithms are given in Fig.6–10, and the adjacent image on the right is a detailed view of the bounded regions.

In Fig. 6, except for PFF-Net and MMP-Net, other models still retain some haze at the intersection of the sky, and MSCNN method produces color distortion. PFF-Net shows grid-like noise, and the proposed MMP-Net performs the best. In Fig. 7, after using dehazing models the texts can be clearly seen in the enlarged picture. Among them, DCP is

TABLE 5. NIQE comparison for dehazed results on real-world images.

image	DCP[9]	CAP[10]	DehazeNet[14]	MSCNN[15]	AOD-Net[16]	GFN[17]	PFF-Net[21]	MMP-Net
Fig.6	3.2299	3.0353	3.3023	3.4812	3.4948	3.0696	4.5930	**2.9761**
Fig.7	3.8035	3.3734	3.5461	3.3290	3.4410	**3.2765**	4.8039	3.5335
Fig.8	3.0860	3.3957	3.2801	3.2923	**3.0749**	3.3745	4.7071	3.2008
Fig.9	4.1899	4.3176	4.3434	4.4294	3.9874	**3.7670**	4.4307	3.9670
Fig.10	1.8721	1.9342	1.9555	1.7793	**1.6950**	2.0685	5.4370	1.8157
Average	3.2363	3.2112	3.2855	3.26	3.1386	3.1112	4.7943	**3.0986**

TABLE 6. Results of different variants of MMP-Net on SOTS.

Model	PSNR	SSIM
MMP-Net without residual block	20.46	0.8267
Proposed MMP-Net	**25.30**	**0.9376**

the clearest, GFN and MMP-Net show the second. However, in terms of the overall subjective perception, DCP is severely distorted in the sky area, while GFN and MMP-Net are indistinguishable. In Fig. 8, it is obvious that in MMP-Net's result the details of the girl's hair can be clearly shown. DCP, GFN and PFF-Net are overexposed, and haze persists in remaining models' results. In Fig. 9, GFN gets the best overall dehazing performance. However, looking at the enlarged detail, only in MMP-Net' result the texture of the eaves are efficiently recovered. MMP-Net also shows superior to other remaining models in terms of overall image dehazing effect. In Fig.10, MSCNN and GFN produce distortion, and CAP's result is excessively dim. Since haze is thin in Fig.10, the results of MMP-Net and other most of models show almost subjective perception. Comprehensive comparison by dehazing performances of 5 real-world hazy images, MMP-Net show the best subjective visual quality both in image details and color fidelity.

Here we use completely blind image quality assessment metric NIQE [27] to verify the effectiveness of MMP-Net in the real-world hazy images from the objective perspective. Notably, the score of NIQE are range from 0 (the best) to 100 (the worst), and the lower NIQE, the better quality. NIQE comparison for dehazed results on real-world images are shown in Table 5. It can be seen proposed MMP-Net has the lowest NIQE score on the average results, suggesting the best dehazing performances.

C. MMP-Net VARIANT EXPERIMENT

A stripping experiment is performed to prove the rationality of the MMP-NET, which includes MMP-Net removing parallel fusion module (MMP-NoPF) and MMP-Net. The MMP-NoPF includes a multi-scale feature extraction module and 18 residual blocks. A multi-scale CNN module can obtain richer image features from the whole to local, such as the contour information and detailed features. The proposed MMP-Net fuses multi-scale features in parallel to improve the robustness and effectiveness of the internal information of the image. The comparisons of PSNR and SSIM for the two models on the SOTS dataset are shown in Fig. 11.

Obviously, MMP-Net with multi-scale features in parallel fusion has the highest PSNR and SSIM, which means it has the best dehazing performance.

D. EFFECTIVENESS OF RESIDUAL BLOCKS

Here, we remove the residual block from MMP-Net to verify the value of residual learning module, and give the results of MMP-Net and MMP-Net without residual block on the SOTS in Table 6.

From Table 6 it can be seen that MMP-Net has obvious advantage in PSNR and SSIM over the MMP-Net removing residual block, which indicates that residual learning module are very useful for extracting image details.

E. RUNNING TIME COMPARISON

In the section, random 50 images in SOTS were selected to test different models, on the same machine (Inter(R) Core(TM) i7-6900K CPU @ 3.20 GHz and Nvidia GeForce GTX 1080 Ti GPUs). The comparision results are given as Table 7 with respect to parameters, model size and runtime. Among them, the runtime refers to the average runtime of the selected 50 images. Obviously, compared with lightweight networks (such as AOD-Net), MMP-Net has more parameters and longer running time. However, compared to deep PFF-Net, MMP-Net consumes shorter runtime and has less parameter.

TABLE 7. Comparison of MMP-Net and state-of-the-art methods with respect to parameters, model size and runtime.

Model	Image size	Param.	Size	Platform	Time(second)
DCP[9]	270x360	-	-	Matlab(C)	7.7293
DehazeNet[14]	270x360	8,240	-	Matlab(C)	1.7194
MSCNN[15]	270x360	8014	-	Matlab(C)	2.2138
AOD-Net[16]	270x360	1761	9Kb	Pytorch (C/G)	0.2696/0.0228
PFF-Net[21]	270x360	14,952,295	57.05Mb	Pytorch (C/G)	0.6543/0.0874
Proposed MMP-Net	270x360	**5,335,060**	**20.37Mb**	**Pytorch(C/G)**	**0.5163/0.0384**

V. CONCLUSION

In this paper, an end-to-end multi-scale feature multiple parallel fusion network MMP-Net is proposed to tackle the challenging single image dehazing problem. Different scale features are extracted by a multi-scale CNN module, and fused them multiple times in parallel. Detailed features are learned by residual block, and are merged with these features from CNNs in parallel. Finally, these features from MMP-Net are utilized to obtain the dehazed image by using the atmospheric scattering model. The experimental results show that our proposed MMP-Net gets the highest PSNR and SSIM among start-of-the-art dehazing methods on the average of the three datasets SOTS, HSTS and D-Hazy, and also gains the best subjective visual quality on real-world hazy images.

REFERENCES

[1] K. B. Gibson, D. T. Vo, and T. Q. Nguyen, "An investigation of dehazing effects on image and video coding," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 662–673, Feb. 2012.
[2] H. Zhang, J. Yang, Y. Zhang, N. M. Nasrabadi, and T. S. Huang, "Close the loop: Joint blind image restoration and recognition with sparse representation prior," in *Proc. IEEE Int. Conf. Comput. Vis.*, Nov. 2011, pp. 770–777.
[3] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.
[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
[5] T.-H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma, "PCANet: A simple deep learning baseline for image classification?" *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5017–5032, Dec. 2015.
[6] Y. Xu, J. Wen, L. Fei, and Z. Zhang, "Review of video and image defogging algorithms and related studies on image restoration and enhancement," *IEEE Access*, vol. 4, pp. 165–188, 2016.
[7] X. Hu, L. Zhuo, and X. Li, "A moving objects based real-time defogging method for traffic monitoring videos," in *Proc. IEEE Int. Conf. Digit. Signal Process.*, Aug. 2014, pp. 1–6.
[8] R. T. Tan, "Visibility in bad weather from a single image," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Aug. 2008, pp. 1–8.
[9] 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.
[10] 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.
[11] D. Berman and S. Avidan, "Non-local image dehazing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Dec. 2016, pp. 1674–1682.
[12] L. Kratz and K. Nishino, "Factorizing scene albedo and depth from a single foggy image," in *Proc. IEEE Conf. Comput. Vis.*, Sep. 2009, pp. 1701–1708.
[13] J. Zhang, Y. Cao, S. Fang, Y. Kang, and C. W. Chen, "Fast haze removal for nighttime image using maximum reflectance prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 7016–7024.
[14] 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.
[15] 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.*, Sep. 2016, pp. 154–169.
[16] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "AOD-Net: All-in-one dehazing network," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 4770–4778.
[17] H. Zhang and V. M. Patel, "Densely connected pyramid dehazing network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3194–3203.
[18] D. Engin, A. Genc, and H. K. Ekenel, "Cycle-dehaze: Enhanced cyclegan for single image dehazing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 3, Jun. 2018, pp. 825–833.
[19] J. Zhang, Y. Cao, Y. Wang, C. Wen, and C. Chen, "Fully point-wise convolutional neural network for modeling statistical regularities in natural images," in *Proc. ACM Multimedia Conf. Multimedia Conf.*, 2018, pp. 984–992.
[20] Y. Du and X. Li, "Recursive deep residual learning for single image dehazing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 730–737.
[21] K. Mei, A. Jiang, J. Li, and M. Wang, "Progressive feature fusion network for realistic image dehazing," in *Proc. Asian Conf. Comput. Vis.*, 2018, pp. 203–215.
[22] 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, Jul. 2019.
[23] A. Dudhane and S. Murala, "RYF-Net: Deep fusion network for single image haze removal," *IEEE Trans. Image Process.*, vol. 29, pp. 628–640, 2019.
[24] S. Kuanar, K. R. Rao, D. Mahapatra, and M. Bilas. (2019). *Night Time Haze and Glow Removal using Deep Dilated Convolutional Network*. [Online]. Available: https://arxiv.xilesou.top/abs/1902.00855
[25] 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 Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3253–3261.
[26] A. Wang, W. Wang, J. Liu, and N. Gu, "AIPNet: Image-to-image single image dehazing with atmospheric illumination prior," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 381–393, Jan. 2019.
[27] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer," *IEEE Signal Process. Lett.*, vol. 22, no. 3, pp. 209–212, Mar. 2013.
[28] E. J. McCartney, *Optics of the Atmosphere: Scattering by Molecules and Particles*. New York, NY, USA: Wiley, 1976, p. 1976.
[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
[30] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 1026–1034, Dec. 2015.
[31] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from RGBD images," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2012, pp. 746–760.
[32] 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.
[33] C. Ancuti, C. O. Ancuti, and C. De Vleeschouwer, "D-Hazy: A dataset to evaluate quantitatively dehazing algorithms," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2016, pp. 2226–2230.
[34] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in PyTorch," in *Proc. 31st Conf. Neural Inf. Process. Syst.*, Long Beach, CA, USA, 2017, pp. 1–4.
[35] Y. Wang, J. Li, Y. Lu, Y. Fu, and Q. Jiang, "Image quality evaluation based on image weighted separating block peak signal to noise ratio," in *Proc. Int. Conf. Neural Netw. Signal Process.*, vol. 2, pp. 994–997, Apr. 2003.
[36] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

JIAJIA YAN was born in Nanjing, Jiangsu, in 1996. She received the bachelor's degree in computer science and technology from Huaiyin Normal University, China, in 2018. She is currently pursuing the master's degree with the Institute of Logistics Science and Engineering Research, Shanghai Maritime University. Her current research direction is image processing.

CHAOFENG LI (Senior Member, IEEE) received the B.S. and M.S. degrees in mathematical geology and the Ph.D. degree in remote sensing image processing from the Chinese University of Mining and Technology, Xuzhou, China, in 1995, 1998, and 2001, respectively.

From July 2001 to June 2003, he has been finishing postdoctoral research work in pattern recognition and intelligent systems at the Nanjing University of Science and Technology. He was a Visiting Research Scholar with The University of Texas at Austin, from May 2008 to May 2009. He is currently an Associate Professor and a Professor with the School of Internet of Things Engineering, Jiangnan University, from May 2003 to December 2017. He is also a Professor with the Institute of Logistics Science and Engineering, Shanghai Maritime University. He has published over 80 articles in reputable journals and conference proceedings, covering the areas of neural networks and image processing and analysis. His current research interests include image processing, machine learning, and medical image analysis.

YUHUI ZHENG was born in China, in 1982. He received the B.S. degree in chemistry and the Ph.D. degree in computer science from the Nanjing University of Science and Technology, Nanjing, China, in 2004 and 2009, respectively. From 2014 to 2015, he was a Visiting Scholar with the Digital Media Laboratory, School of Electronic and Electrical Engineering, Sungkyunkwan University, South Korea. He is currently an Associate Professor with the School of Computer and Software, Nanjing University of Information Science and Technology. His research interests include image processing, pattern recognition, and remote sensing information systems.

SHOUKUN XU was born in China, in 1972. He received the B.S., M.S., and Ph.D. degrees from the Chinese University of Mining and Technology, Xuzhou, China, in 1995, 1998, and 2001, respectively. He is currently a Professor with the School of Information Science and Engineering, Changzhou University. His research interests include computer vision, soft development, and computer education. He has gained the two second prizes of Scientific and Technological Progress Award of Provincial/Ministerial Level, and one Second Prize of National Teaching Achievement Award, the Award of Outstanding Young Teacher (Qinglan Project) of Jiangsu Province.

XIAOYONG YAN received the Ph.D. degree in computer science and technology from the Nanjing University of Science and Technology, in 2013. He is currently an Associate Professor with the Nanjing University of Posts and Telecommunications, Nanjing, China. His research interests are machine learning, pattern recognition, and the Internet of Things.