Bidirectional Deep Residual learning for Haze Removal

Guisik Kim, Jinhee Park, Suhyeon Ha, and Junseok Kwon
School of Computer Science and Engineering, Chung-Ang University, Seoul, Korea
specialre@naver.com   iv4084em@gmail.com   tngus3752@gmail.com   jskwon@cau.ac.kr

Abstract

Recently, low-level vision problems has been addressed using residual learning that can learn a discrepancy between hazy and haze-free images. Following this approach, in this paper, we present a new dehazing method based on the proposed bidirectional residual learning. Our method is implemented by generative adversarial networks (GANs), consisting of two components, namely, haze-removal and haze-reconstruction passes. The method alternates between removal and reconstruction of hazy regions using the residual to produce more accurate haze-free images. For efficient training, we adopt a feature fusion strategy based on extended tree-structures to include more spatial information and apply spectral normalization techniques to GANs. The effectiveness of our method is empirically demonstrated by quantitative and qualitative experiments, indicating that our method outperforms recent state-of-the-art dehazing algorithms. In particular, our approach can be easily used to solve other low-level vision problems such as deraining.
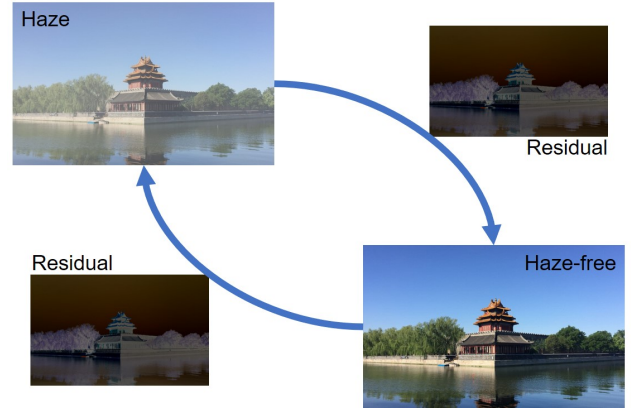
Figure 1. **Illustration of the proposed dehazing method incorporating our residual learning.** Residual refers to the difference between hazy and haze-free images. Thus, a haze-free image can be obtained by subtracting the residual from the haze image. Conversely, addition of the residual to a haze-free image can produce a haze image. Our method aims to generate realistic hazy and haze-free images by learning them in an alternating manner.

1. Introduction

Owing to the emergence of deep learning, high-level vision methods have made a remarkable progress. However, they still suffer from many obstacles contained in real-world images (*e.g.* the shortage of light and bad weather such as haze, rain, night), leading to worse results in the real environments than those in ideal environments. In particular, a haze extremely reduces visibility of images, resulting in the degradation of the performance of most vision methods. In this context, dehazing is being very important as a preprocessor of high-level vision tasks. There are two non-deep-learning approaches to dehazing. One is to utilize additional information such as 3d model, numerous sensors, and multiple images for dehazing [23, 34, 38]. The other is based on only a single image which has been actively studied, showing its effectiveness. The most important factor in dehazing is the estimation of transmission and atmospheric light, making the performance of traditional dehazing methods to highly depend on the accuracy of the estimation. While

end-to-end deep learning methods can alleviate this problem, they typically require haze-free images as ground-truth that are very difficult to acquire in real-world. To solve this problem, ground-truth has been synthesized by generating transmission either using the haze model or using depth information, where depth information can be obtained by single image depth estimation techniques or specific sensors. With these synthesized training data, deep learning methods have shown remarkable improvements in the application of dehazing.

Recently, generative adversarial network (GANs) have driven advanced performance in a variety of applications of computer visions such as image generation [27, 35], style transfer [7, 18, 41], segmentation [11], and object detection [42]. In particular, successful applications of GANs to dehazing have been also presented. The optimization in the GAN applications often involve residual learning that can alleviate gradient vanishing problems, in which gradients converge to zeros as the network becomes deeper. Residual is indicative of fitness between the input image and ground

truth. In the application of dehazing, residual can be considered as differences between hazy and hazy-free images. Thus, by using simple arithmetic operations (*e.g.* sum and subtraction of residuals), we can generate hazy and hazy-free images in an alternating manner, as shown in Fig.1.

In this paper, we propose a GAN-based dehazing method that considers residuals in the generation of hazy and haze-free images. Although residual learning is a well-known technique, using the residual for dehazing has not been actively studied. In contrast to traditional dehazing methods, our method directly applies residuals to images in a bidirectional way to generate both hazy and haze-free images, leading to more accurate results. For this bidirectional generation, we adopt a deep neural architecture similar to CycleGAN [48]. Fig.2 shows the overall procedure of our method.

The contributions of our method are threefold.

• We propose a bidirectional residual learning method for dehazing, in which the haze-removal pass network generates haze-free images from hazy images and the haze-reconstruction pass network recovers hazy images from generated haze-free images.

• We present a feature fusion method for efficient residual learning, which is based on the extended tree-structure.

• We demonstrate the effectiveness and applicability of our bidirectional residual learning method by using the method to solve other low-level vision tasks such as deraining.

In the remainder of this paper, we relate our method to the prior works in Sec. 2 and propose the bidirectional residual learning method in Sec. 3. The implementation of our bidirectional residual learning is presented in Sec. 4.1. Experimental results are provided in Sec. 4 and the conclusion follows in Sec. 5.

2. Related Work

While there have been a variety of algorithms that use additional information for dehazing [23, 34, 38], we discuss in more detail the closely related works to our method, in which only a single image is used for dehazing (*e.g.* dark channel prior [16]).

Koschieder*et al.* [24] defined the haze model as:

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

where I is the haze image, J is the haze-free image, A is the atmospheric light, and t is the transmission. The transmission map keeps depth information of the image, thus atmospheric scattering can be represented as:

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

where β is a scattering parameter and d is depth. Color-line [13] and Haze-line [4] gained impressive dehazing results, especially in color restoration and accuracy. Kim *et al.* estimated atmospheric light effectively with quad-tree searching [22]. Ancuti *et al.* conducted haze region detection with the semi-inverse method and applied multi scale fusion [2]. Additionally, they proposed another scheme that employs multiple features [3]. Kim and Kwon made use of the illumination map as pixel-wise atmospheric light based on retinex theory. Meng *et al.* suggested boundary constant and contextual regularization methods [30]. Choi *et al.* presented the referenceless evaluation method for dehazing in which ground-truth is not required [8]. Zhu *et al.* proposed the color attenuation prior method [49]. To restrain the halo effect (*i.e.* remained haze after dehazing), Chen *et al.* attempted to reduce various artifacts with gradient residual minimization [6].

Deep neural networks for transmission and atmospheric light estimation [28, 36, 5] have emerged as a way to dehaze images. DehazeNet [5] was the first end-to-end dehazing algorithm based on the regression network. Ren *et al.* developed the multi-scale convolutional neural network (MSCNN) [36] via a coarse-to-fine approach for dehazing and introduced the fusion-based method that uses multiple inputs: white balancing, gamma correction, and contrast enhancing [37]. Zhang and Patel removed haze by the dense connection and pyramid method [47]. AOD-Net [25] reformulated the atmospheric scattering model for the end-to-end dehazing algorithm and showed that joint learning of detection and dehazing can improve the performance of object detection. A successful example of image-to-image translation methods [47, 46] for deahzing is Cycle-Dehaze [12] that adopted the CycleGAN architecture in [48] to dehazing. Du and Li used residual learning [10] for dehazing, which is generally used in super resolution problems. They presented POGAN [9] combined with adversarial learning. A advantage of residual learning is that the estimation of transmission and atmospheric light is not required and the haze model is not necessary in the course of dehazing. Thus, it can simplify the deep neural architecture and utilize a variety of existing techniques.

We consider POGAN [9] as a baseline. Unlike POGAN, we present the bidirectional residual learning method for dehazing, adopting both residual learning (*e.g.* ResNet [17]) and consistency learning (*e.g.* CycleGAN [48]) to remove haze.

3. Residual Learning for Dehazing

Residual learning has been used widely in computer vision problems with deep learning [21, 10, 9, 14]. It has been adopted first to super resolution problems but other problems in computer vision have started to be considered as a form: $Y = X - R$. Y is the clear image that we want to obtain, while X is the image with haze, rain, or snow. We can acquire the clear image by removing the residual (*e.g.* haze, rain or snow). In particular, this idea has mostly been used
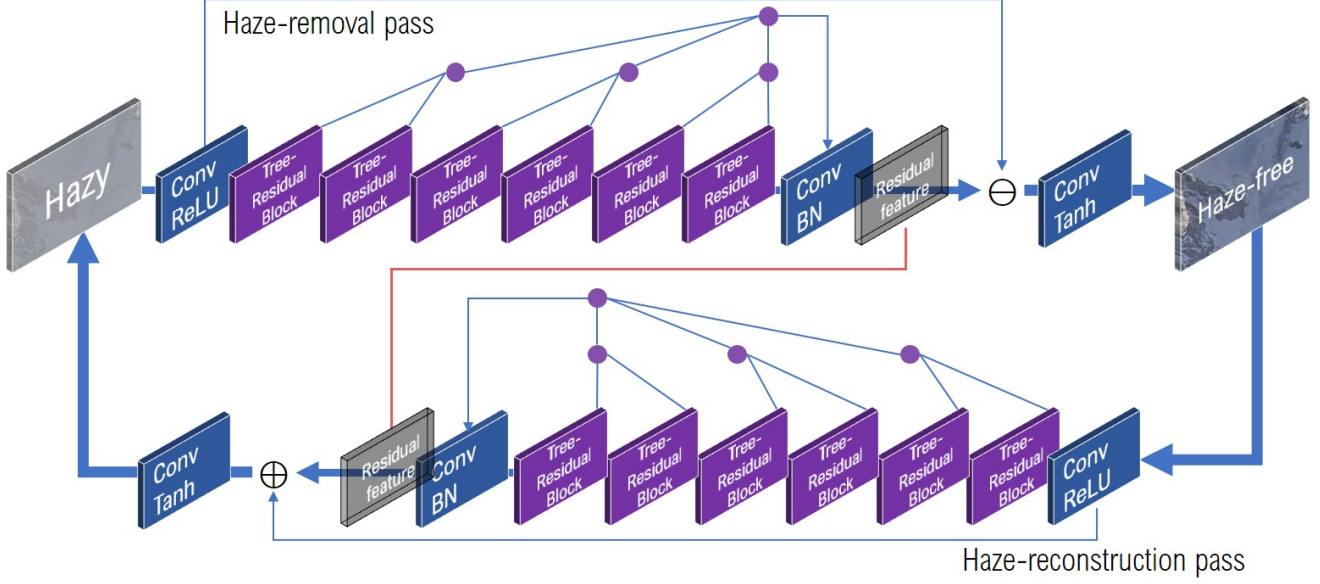
Figure 2. **Overall architecture of the proposed method**.

in deraining [9, 44, 45], because rain streak fits with the residual learning setting. The method in [10] firstly considers haze as residual, implying that the estimation of transmission and atmospheric light is not required in the course of dehazing.

For dehazing, residual learning is reformulated as follows:

$$J(x) = I(x) - R(x), \tag{3}$$

where $J(x)$ is the haze-free image, $I(x)$ is the hazy image, and $R(x)$ is residual image. The residual image is represented as the discrepancy map between haze-free and haze images. Then, the loss function for residual learning in dehazing can be designed as:

$$\mathcal{L}(\theta) = \|F(I(x)) - (I(x) - J(x))\|^2, \tag{4}$$

where θ indicate parameters to train and $F(\cdot)$ is a generator network. With the loss function in (4), we can train the deep neural network for dehazing because the estimated residual can be used to obtain the haze-free image by subtracting residual from the hazy image. While traditional convolutional neural networks has several limits in the estimation of residual for dehazing, we introduce the GAN based residual learning method for dehazing.

3.1. Bidirectional Residual Learning using GAN

In this paper, we propose a bidirectional residual learning method that can be applicable for both haze removal and haze reconstruction. In our method, residual features are obtained in bidirectional ways and two different arithmetic operations are applied to residual learning. We assume that

residual features to obtain haze-free images should be same as those to obtain hazy images.

For bidirectional residual learning, we define two directions, namely, haze-removal pass $G : X-> Y$ and haze-reconstruction pass $F : Y-> X$, respectively. X and Y are hazy and haze-free images, respectively. G and F are generator networks for haze-removal and haze-reconstruction passes, respectively. $R(x)$ and $R(x)'$ are residual features for haze-removal and haze-reconstruction passes, respectively. Then, the adversarial loss function for the haze-removal pass is as follows:

$$\mathcal{L}_{GAN} = \sum_{n=1}^{N} -D_X\big(G(I(x))\big) + \sum_{n=1}^{N} -D_Y\big(F(G(I(x)))\big), \tag{5}$$

where $D(\cdot)$ is a discriminator network. For stable training, we use WGAN-GP [15] and spectral normalization [33].

Our method also uses the content loss function that measures the the discrepancy between two images at the feature level via the l_2 norm, as follows:

$$\mathcal{L}_{content} = \|\phi(G(I(x))) - \phi(J(x))\|^2 + \|\phi(F(G(I(x)))) - \phi(I(x))\|^2, \tag{6}$$

where ϕ is the 51st layer of the VGG network [39], which is a more deeper layer than that used in the perceptual loss [20]. The feature map has a lot of information, leading to slow learning in some cases. Thus, we use the deeper layer with prominent features and obtain the content loss on the high level feature map. Note that the contents loss in (6) prevents the generation of blurry images by imposing more powerful constraints compared with pixel level losses.
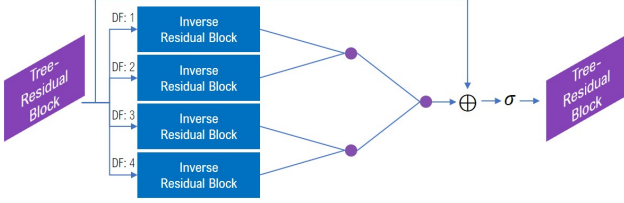
Figure 3. **Our feature fusion method** based on extended tree-structure and inverse residual blocks.

(a) Haze removal   (b) Haze reconstruction

Figure 4. **Residual feature maps** estimated by haze-removal pass network and haze-reconstruction pass network on buildings image.

For spatial smoothness, the total variation loss function is defined as follows:

$$\mathcal{L}_{tv} = \| \bigtriangledown_x G(I(x)) + \bigtriangledown_y G(I(x)) \|^2 + \\ \| \bigtriangledown_x F(G(I(x))) + \bigtriangledown_y F(G(I(x))) \|^2, \quad (7)$$

where $\bigtriangledown_x$ and $\bigtriangledown_y$ denote gradients in x and y directions, respectively. The following function calculates the L1 loss between generator outputs and ground-truth:

$$\mathcal{L}_{l1} = \| G(I(x)) - J(x) \|_1 + \| F(G(I(x))) - I(x) \|_1. \quad (8)$$

While conventional L2 loss functions induce blurry images, the L1 loss function in (8) can produce clearer images. Before passing through the final $tanh$ layer, the haze-removal pass network applies element-wise subtraction, whereas the haze-reconstruction pass network applies element-wise summation. Since the residual features obtained by two networks must be equal to each other, we add the consistency loss function $\mathcal{L}_{consistency}$ for residual features, $R(x)$ and $R(x)'$:

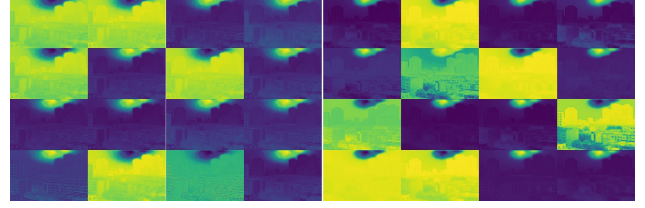$$\mathcal{L}_{consistency} = \| R(x) - R(x)' \|_1. \quad (9)$$

The total loss function for the proposed bidirectional residual learning is defined as follows:

$$\mathcal{L}_{total} = \mathcal{L}_{GAN} + \lambda_1 \cdot \mathcal{L}_{content} + \lambda_2 \cdot \mathcal{L}_{SSIM} + \\ \lambda_3 \cdot \mathcal{L}_{l1} + \lambda_4 \cdot \mathcal{L}_{tv} + \lambda_5 \cdot \mathcal{L}_{consistency}, \quad (10)$$

where the SSIM loss function [43] makes structures of the estimated image and ground-truth similar. Using The SSIM loss, our method can preserve local details that can be lost by the content and L1 loss functions. λ_1, λ_2, λ_3, λ_4, and λ_5 in (10) are control parameters, balancing weights of different loss functions. These parameters are empirically determined and fixed to $50, 1, 0.05, 0.1, 0.05$, respectively. Note that the weight ratio of the contents loss to the GAN loss is typically set to $100 : 1$.

3.2. Tree-Residual Block

Feature fusion techniques yield better training results. For example, JORDER *et al*. [45] concatenated multiple features extracted by using different dilated convolutions. Fu *et al*. [14] fused multiple features in the form of a tree-structure. In contrast to the above methods, we present a tree-residual structure that can more accurately preserve spatial information by applying residual blocks to feature extraction. The basic tree-structure formula is as follows:

$$Z = \sigma(W_{fuse} * concat(Z_1, Z_2) + b), \quad (11)$$

where Z_1 and Z_2 are adjacent features and have the same dimension. W_{fuse} combines two feature maps with the 1×1 convolution. After activation, the fused feature Z has the same dimension as original features Z_1 and Z_2. Our method uses residual blocks to extract features such as Z_1 and Z_2 and makes tree structures inside of each residual block and between multiple residual blocks. Fig.3 illustrates the proposed tree-residual structure.

3.3. Network Architecture

The proposed generator network is illustrated in Fig.2. The network contains six tree-residual blocks after the first conv-relu layer and four inverse residual blocks for each tree-residual block. For short-cut connection of a residual block, element-wise summation is replaced by element-wise subtraction [17]. It is experimentally demonstrated that element-wise subtraction is better than element-wise summation operation for our residual learning. The tree-residual block uses dilated convolutions, in which the dilated factor is set to $1 \sim 4$ to consider various receptive fields. In tree-residual blocks and inverse residual blocks, features are fused by tree-structures, as explained in Sec. 3.2. Before the last conv-tanh layer, we perform the element-wise subtraction (summation) of the residual feature map and the first convolution layer for the haze-removal (haze-reconstruction) pass network. With these operations, our network can produce both the hazy and haze-free images. We use 16 channels for each feature map.

The discriminator network uses the Wasserstein GAN with gradient penalty [15], PatchGAN [19], and spectral normalization [33]. α in LeakyReLU is set to 0.2. Table 1 describes detailed network structures.

Table 1. **Parameter settings on the generator network.** "conv" denotes the convolution layer. Other layers (*e.g.* batch normalization and activation function) are omitted and repeated layers are mentioned only once in the table. The detailed connection for feature fusion is illustrated in Fig 2.

		residual block	residual block	residual block	residual block	Feature fusion	Residual feature	Reconstruction
Layer	conv	conv	conv	conv	conv	conv	conv	conv
Kernel size	3x3	3x3	3x3	3x3	3x3	1x1	3x3	1x1
Stride	1	1	1	1	1	1	1	1
Pad	1	1	2	3	4	0	1	0
Channel	16	16	16	16	16	16	16	3
Dilated factor	1	1	2	3	4	1	1	1

Figure 5. **Intermediate haze-removed** $G(I(X))$ **and haze-reconstructed** $F(G(I(X)))$ **images** generated in the course of our bidirectional residual learning.

4. Experimental Results

4.1. Implementation Details

Training dataset: It is difficult to make ground-truth for real-world hazy images. For indoor images, we can extract the depth maps of images via depth sensors and use these maps as ground-truth for hazy regions. However, these maps typically are dissimilar to real hazy regions. Recently, depth estimation algorithms [29] have been proposed to synthesize hazy images from haze-free images (*i.e.* ground-truth) based on the haze model, in which the transmission parameter β in (2) ranges 0.5 to 1.5 with the step size of 0.1 and the atmospheric light A in (1) is randomly selected over a range of values from 0.7 to 1.0. For haze synthesis, we used DIV2K dataset [1] . We randomly cropped 100×100 patches using 900 images to make $230,000$ patches. To pre-

Table 2. **Quantitative evaluation of reconstructed haze images** on real-world dataset.

	house	mountain	pumpkins	florence	ny17
SSIM	0.909	0.948	0.946	0.877	0.88

vent haze-free regions from being saturated, we applied 1.1 gamma correction to 40 percent of the amount of total images.

Parameter Setting: Patches were resized to 70×70. Mini-batch size was set to 8. We used the loss function in (10) and trained the generator network every five times per iteration. The initial learning rate was set to $1e-3$ and weight decay was 0.05. We used the ADAM optimizer and performed a total of 50k iteration. Training was performed using NVIDIA GeForce GTX TITAN XP GPU and the proposed network was implemented in PyTorch on windows environment.

Comparison: Our method was compared with state-of-the-art dehazing methods [36, 5, 25] based on deep learning and non-deep learning methods [13]. We used real-world datasets such as Fattal [13] and synthesized datasets presented in [26] for qualitative evaluation. We utilized three image quality evaluation methods (*i.e.* naturalness image quality evaluator (NIQE) [32], blind/referenceless image spatial quality evaluator (BRISQUE) [31], and perception based image quality evaluator (PIQE) [40]), which are no-reference quantitative evaluation metrics.

4.2. Ablation Study

Fig.4 shows the residual maps produced by haze-removal pass network and haze-reconstruction pass network, demonstrating that hazy regions were accurately estimated. Fig.5 includes haze-removed and haze-reconstructed images that were produced in the course of bidirectional residual training. Our bidirectional residual training method alternated between removal and restoration of haze in the determination of hazy regions and generated accurate haze-free images. Table 2 reports the SSIM scores on test images used in Fig.5. Most results are close to 0.9 that is close to those of real-world haze images.
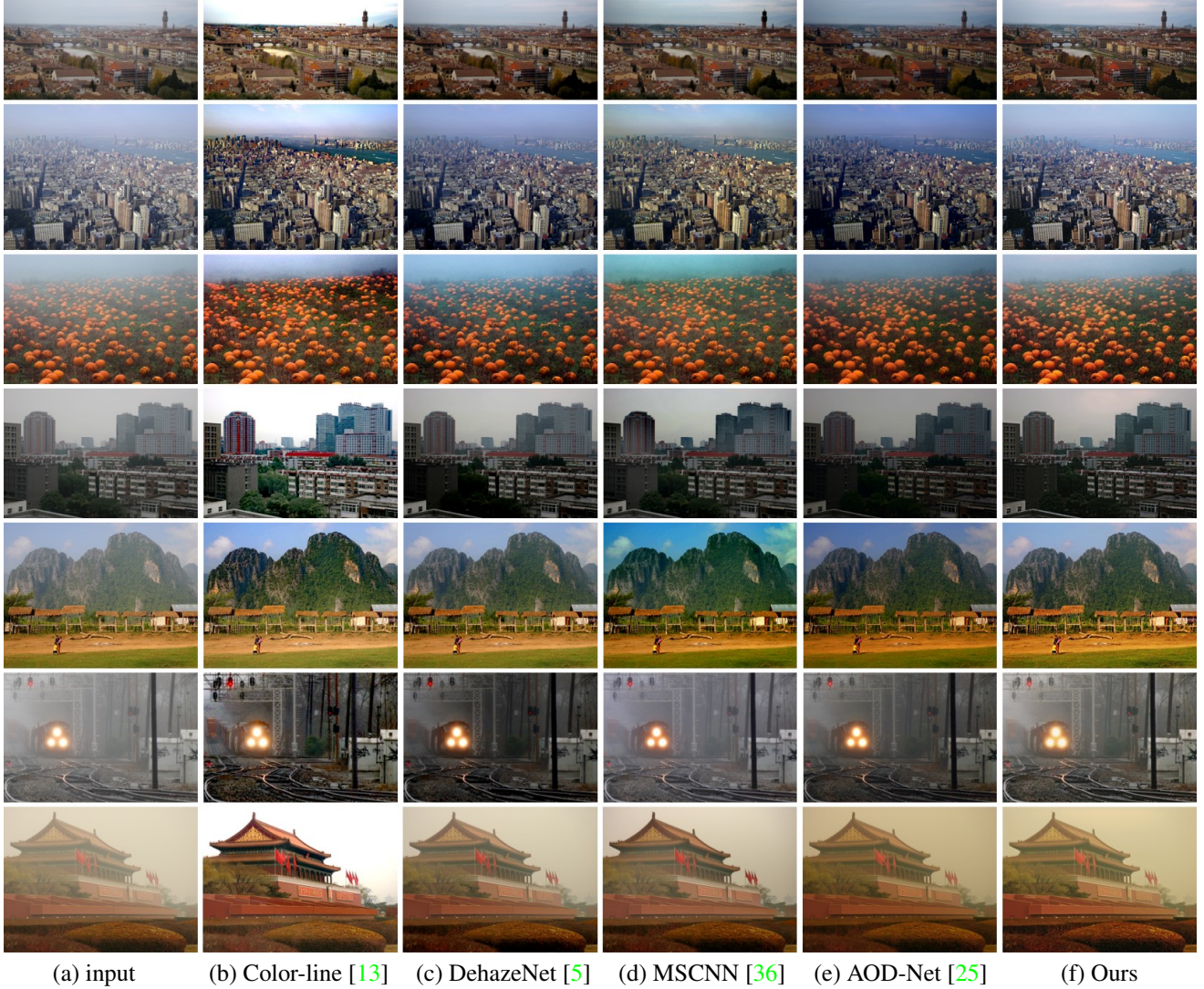
(a) input   (b) Color-line [13]   (c) DehazeNet [5]   (d) MSCNN [36]   (e) AOD-Net [25]   (f) Ours

Figure 6. **Comparison of dehazing results on real-world images.**

Table 3. **Quantitative comparison with sate-of-the-art dehazing methods** using no-reference image quality measurements on real-world images. Red and blue numbers denote the best and the second best results, respectively.

	Color-line [13]	DehazeNet [5]	MSCNN [36]	AOD-Net [25]	Ours
NIQE [32]	3.031	3.238	2.897	3.334	2.967
BRISQUE [31]	18.90	24.75	17.35	17.89	16.40
PIQE [40]	33.23	35.98	30.67	33.23	29.89

4.3. Results on Real-world datasets

We qualitatively compared several dehazing methods with real-world dataset, Fattal [13], as shown in Fig.6. While the color-line method produced more colorful results than other methods, it incurred saturation problems and artefacts as shown in the river of the first row and left-hand building of the fourth row, respectively. DehazeNet, MSCNN, and AOD-Net generally produced accurate de-hazing results, but generated slightly faded and dark tone images,as shown in the fourth row. The proposed method accurately got rid of hazy regions, while images were darkening slightly as the dehazing process proceeded/ However, the method still preserved a lot of details.

Table 3 includes numerical results of state-of-the-art dehazing methods evaluated in term of no-reference image quality metrics with test images of Fig 6. Our method and MSCNN show good performance in general.
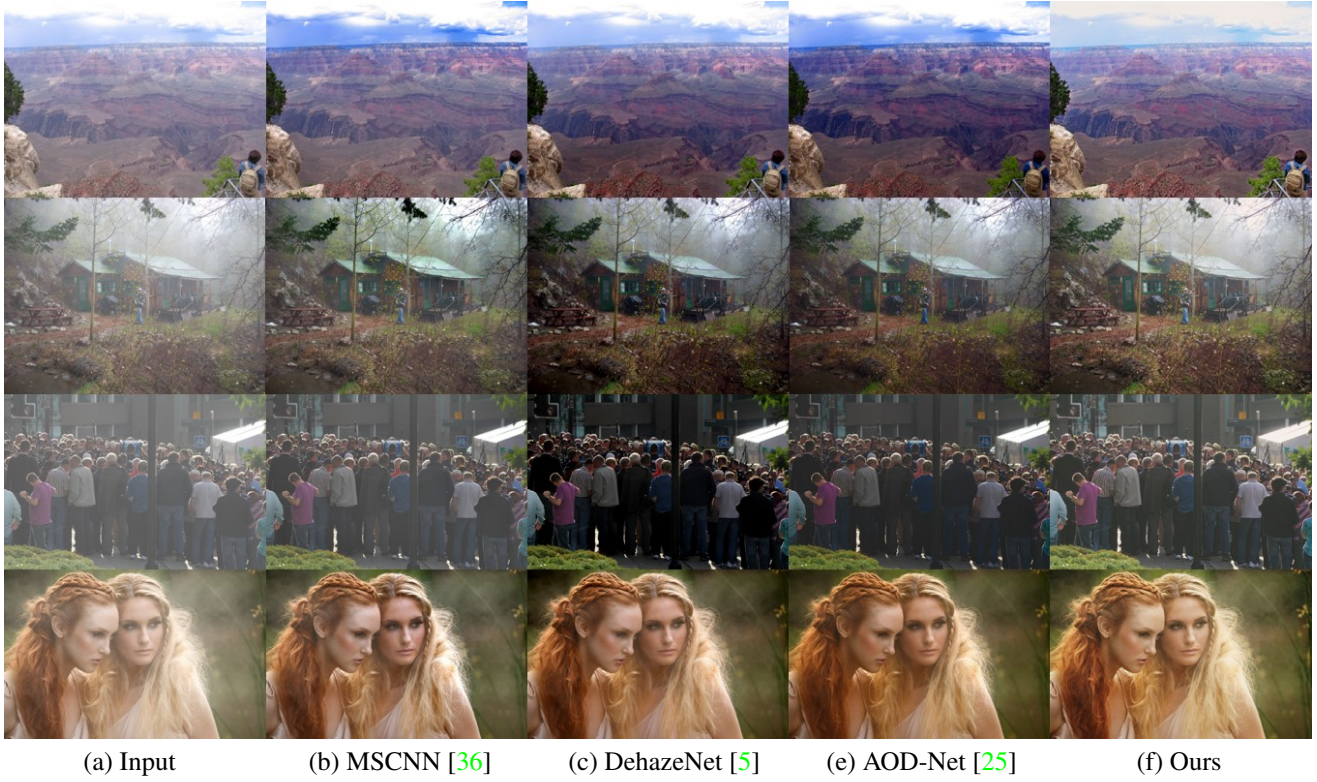
| (a) Input | (b) MSCNN [36] | (c) DehazeNet [5] | (e) AOD-Net [25] | (f) Ours |

Figure 7. **Comparison of dehazing results on real-world images.**

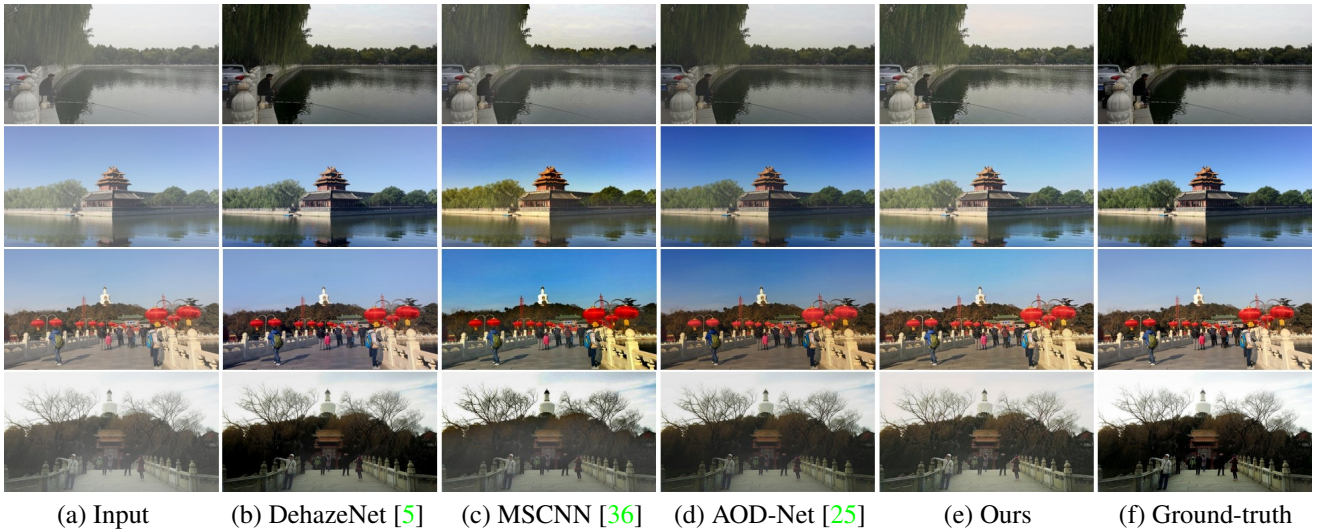| (a) Input | (b) DehazeNet [5] | (c) MSCNN [36] | (d) AOD-Net [25] | (e) Ours | (f) Ground-truth |

Figure 8. **Comparison of dehazing results on synthetic images.**

4.4. Results on Synthetic datasets

RESIDE [26] presented a variety of synthesized datasets for dehazing. Since indoor images were not suitable to describe real-world environments, outdoor images that contain heavy haze were used for evaluation. Fig.8 shows dehazing results of several methods. While MSCNN removed only a part of hazy regions, AOD-Net decreased the bright-ness of sky areas. In contrast to the above methods, our method and DehazeNet produced the most natural results.

4.5. Deraining results

A rain streak is also residual that can be obtained by measuring a discrepancy between rain and clear images. Thus, for deraining, the residual learning based methods

(a) Rain images   (b) Our deraining results

Figure 9. **Results of deraining using proposed residual learning.**

have showed good performance [45, 14]. Our bidirectional residual learning method can be also adopted easily to solve this deraining problems. For evaluation for deraining, We used the Rain100H dataset [45]. Fig.9 shows the deraining results. Our method accurately removed the rain streak, while making other regions clear.

5. Conclusion

In this paper, we proposed a dehazing method using bidirectional residual learning. Our method considered residual that is a difference between haze images and haze-free images and alternated between removal and reconstruction of hazy regions using the residual. For alternative generation of hazy and haze-free images, we present the haze-removal pass network and the haze-reconstruction network. This alternation process helped to improve the quality of haze-free images. Additionally, the feature fusion method based on the extended tree-structure was proposed, which enables our method to use more accurate spatial information. Experimental results demonstrated that Our method outperforms other methods for dehazing. Our method can be easily adopted to solve various low-level vision such as deraining.

6. Acknowledgements

References

[1] E. Agustsson and R. Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *CVPR Workshops*, 2017. 5

[2] C. Ancuti, C. Ancuti, C. Hermans, and P. Bekaert. A fast semi-inverse approach to detect and remove the haze from a single image. In *ACCV*, 2011. 2

[3] C. O. Ancuti and C. Ancuti. Single image dehazing by multi-scale fusion. *TIP*, 22(8):3271–3282, 2013. 2

[4] D. Berman, T. Treibitz, and S. Avidan. Non-local image dehazing. In *CVPR*, 2016. 2

[5] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao. DehazeNet: An end-to-end system for single image haze removal. *TIP*, 25(11):5187–5198, 2016. 2, 5, 6, 7

[6] C. Chen, M. N. Do, and J. Wang. Robust image and video dehazing with visual artifact suppression via gradient residual minimization. In *ECCV*, 2016. 2

[7] Y. Chen, Y.-K. Lai, and Y.-J. Liu. Cartoongan: Generative adversarial networks for photo cartoonization. In *CVPR*, 2018. 1

[8] L. K. Choi, J. You, and A. C. Bovik. Referenceless prediction of perceptual fog density and perceptual image defogging. *TIP*, 24(11):3888–3901, 2015. 2

[9] Y. Du and X. Li. Perceptually optimized generative adversarial network for single image dehazing. *arXiv preprint arXiv:1805.01084*, 2018. 2, 3

[10] Y. Du and X. Li. Recursive deep residual learning for single image dehazing. In *CVPR Workshops*, 2018. 2, 3

[11] K. Ehsani, R. Mottaghi, and A. Farhadi. Segan: Segmenting and generating the invisible. In *CVPR*, 2018. 1

[12] D. Engin, A. Genç, and H. Kemal Ekenel. Cycledehaze: Enhanced cyclegan for single image dehazing. In *CVPR Workshops*, 2018. 2

[13] R. Fattal. Dehazing using color-lines. *TOG*, 34(13):13:1–13:14, 2014. 2, 5, 6

[14] X. Fu, Q. Qi, Y. Huang, X. Ding, F. Wu, and J. Paisley. A deep tree-structured fusion model for single image deraining. *arXiv preprint arXiv:1811.08632*, 2018. 2, 4, 8

[15] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville. Improved training of wasserstein gans. In *NIPS*, 2017. 3, 4

[16] K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. In *CVPR*, 2009. 2

[17] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *CVPR*, 2016. 2, 4

[18] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial nets. In *CVPR*, 2017. 1

[19] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In *CVPR*, 2017. 4

[20] J. Johnson, A. Alahi, and L. Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *ECCV*, 2016. 3

[21] J. Kim, J. K. Lee, and K. M. Lee. Accurate image super-resolution using very deep convolutional networks. In *CVPR*, 2016. 2

[22] J.-H. Kim, W.-D. Jang, J.-Y. Sim, and C.-S. Kim. Optimized contrast enhancement for real-time image and video dehazing. *J VIS COMMUN IMAGE R.*, 24(3):410–425, 2013. 2

[23] J. Kopf, B. Neubert, B. Chen, M. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, and D. Lischinski. Deep photo: Model-based photograph enhancement and viewing. *TOG*, 27(5):116, 2008. 1, 2

[24] H. Koschmieder. *Theorie der horizontalen sichtweite: kontrast und sichtweite*. Keim & Nemnich, 1925. 2

[25] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng. AOD-Net: All-in-one dehazing network. In *ICCV*, 2017. 2, 5, 6, 7

[26] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang. Benchmarking single-image dehazing and beyond. *TIP*, 28(1):492–505, 2019. 5, 7

[27] J. Lin, Y. Xia, T. Qin, Z. Chen, and T.-Y. Liu. Conditional image-to-image translation. In *CVPR*, 2018. 1

[28] Z. Ling, G. Fan, J. Gong, and S. Guo. Learning deep transmission network for efficient image dehazing. *Multimed Tools Appl.*, 78(1):213—236, 2019. 2

[29] F. Liu, C. Shen, G. Lin, and I. Reid. Learning depth from single monocular images using deep convolutional neural fields. *TPAMI*, 38(10):2024–2039, 2016. 5

[30] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan. Efficient image dehazing with boundary constraint and contextual regularization. In *ICCV*, 2013. 2

[31] A. Mittal, A. K. Moorthy, and A. C. Bovik. No-reference image quality assessment in the spatial domain. *TIP*, 21(12):4695–4708, 2012. 5, 6

[32] A. Mittal, R. Soundararajan, and A. C. Bovik. Making a "completely blind" image quality analyzer. *SPL*, 20(3):209–212, 2013. 5, 6

[33] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida. Spectral normalization for generative adversarial networks. *arXiv preprint arXiv:1802.05957*, 2018. 3, 4

[34] S. G. Narasimhan and S. K. Nayar. Contrast restoration of weather degraded images. *TPAMI*, 25(6):713–724, 2003. 1, 2

[35] S. Reed, Z. Akata, L. X. Yan, B. Logeswaran, Schiele, and H. Lee. Generative adversarial text to image synthesis. In *ICML*, 2016. 1

[36] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang. Single image dehazing via multi-scale convolutional neural networks. In *ECCV*, 2016. 2, 5, 6, 7

[37] W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W. Liu, and M.-H. Yang. Gated fusion network for single image dehazing. In *CVPR*, 2018. 2

[38] L. Schaul, C. Fredembach, and S. Süsstrunk. Color image dehazing using the near-infrared. In *ICIP*, 2009. 1, 2

[39] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. 3

[40] N. Venkatanath, D. Praneeth, M. C. Bh, S. S. Channappayya, and S. S. Medasani. Blind image quality evaluation using perception based features. In *NCC*, 2015. 5, 6

[41] X. Wang and A. Gupta. Generative image modeling using style and structure adversarial networks. In *ECCV*, 2016. 1

[42] X. Wang, A. Shrivastava, and A. Gupta. A-fast-rcnn: Hard positive generation via adversary for object detection. In *CVPR*, 2017. 1

[43] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, et al. Image quality assessment: from error visibility to structural similarity. *TIP*, 13(4):600–612, 2004. 4

[44] W. Yang, J. Feng, J. Yang, F. Zhao, J. Liu, Z. Guo, and S. Yan. Deep edge guided recurrent residual learning for image super-resolution. *TIP*, 26(12):5895–5907, 2017. 3

[45] W. Yang, R. T. Tan, J. Feng, J. Liu, Z. Guo, and S. Yan. Deep joint rain detection and removal from a single image. In *CVPR*, 2017. 3, 4, 8

[46] X. Yang, Z. Xu, and J. Luo. Towards perceptual image dehazing by physics-based disentanglement and adversarial training. In *AAAI*, 2018. 2

[47] H. Zhang and V. M. Patel. Densely connected pyramid dehazing network. In *CVPR*, 2018. 2

[48] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *ICCV*, 2017. 2

[49] Q. Zhu, J. Mai, and L. Shao. A fast single image haze removal algorithm using color attenuation prior. *TIP*, 24(11):3522–3533, 2015. 2