



A deep hourglass-structured fusion model for efficient single image dehazing

Yufeng Li¹ · Xiang Chen¹ · Caihua Kong¹ · Longgang Dai¹ · Yufeng Huang¹

Received: 25 November 2020 / Revised: 27 August 2021 / Accepted: 17 January 2022 /
Published online: 16 February 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Hazy images obstruct the visibility of image content, which can negatively affect vision-based decision-making in multimedia systems and applications. Recently, convolutional neural networks (CNN) are proven with great benefit to remove single image haze, which has aroused research attention. However, in practice, previous works fail to fully exploit multi-scale features and restore the faithful image details from the hazy inputs, resulting in sub-optimal performance. In this paper, we propose a novel and high-efficiency deep hourglass-structured fusion model to address this issue, which also indicates the applicability of the modified hourglass architecture to remove haze. Unlike the conventional multi-scale learning schemes, top-down and bottom-up feature fusions are repeated, so each of the coarse-to-fine scale representations receives data of parallel ones, which allows for more flexible information exchange and aggregation at various scales. To be specific, we develop residual dense module as the backbone unit, while introducing the channel-wise attention mechanism to further enhance the representation ability of the network. As proved by extensive assessments demonstrate, our designed model out-classes existing ones and achieves the advanced performance on benchmark datasets and real hazy images. We have released source codes on GitHub: <https://github.com/cxtalk/Hourglass-DehazeNet>.

Keywords Single image dehazing · Hourglass architecture · Multi-scale features · Parallel fusion · Channel attention · Deep learning

1 Introduction

Haze refers to a bad weather phenomenon, and images captured in such scenario undergo visible quality degradation (e.g., color distortion and information loss). Such degradation of

✉ Xiang Chen
cv.xchen@gmail.com

images severely hinders the performance of many multimedia applications such as pedestrian detection, autonomous driving, and video surveillance, etc. Accordingly, haze removal has been adopted as a vital preprocessing step and receives significant attention recently in the computer vision field [4, 15, 28, 37].

Referring dehazing process, numerous approaches are carried out for addressing such tough issue. Early works are mostly prior based, which leverage hand-crafted statistics as extra mathematical constraints for remedying data lost in the corrupting process. Here are the following instances. He et al. [10] fully exploited dark channel prior (DCP) to estimate map of transmitting process. Berman et al. [1] developed a non-local prior to observe the formation of compact clusters in RGB space according to the color of clear images. Although these approaches are simple and fast to perform, there are still some limitations in specific cases.

Recently, we have witnessed the rapid progress of convolutional neural networks (CNN) for image dehazing task as supported by the powerful feature representation capability. Earlier dehazing methods used deep neural networks to estimate the transmission map and atmospheric light, including DehazeNet [2], Multi-Scale CNN (MSCNN) [25], All-in-One Dehazing Network (AOD-Net) [14]. Ren et al. [26] exploited the Gated Fusion Network (GFN), which leverages hand-selected preprocessing strategies and multi-scale estimation. Although the above methods perform in certain dehazing tasks, they may not be able to estimate the transmission map by depending on atmospheric scattering model. Subsequently, there are some methods that use end-to-end deep models to directly estimate clean results instead of explicitly estimating the parameters of the atmospheric scattering model. Considering that the estimation of the transmission map and atmospheric light occasionally deviates from the real hazy image, directly estimating the haze-free image can avoid sub-optimal restoration. For instance, the work of [3] offer a Gated Context Aggregation network (GCA-Net) for removing the gridding artifacts attributed to dilated convolution. Unfortunately, it is difficult for such a single-scale framework to capture the intrinsic correlation of haze characteristics at different scales. Later, Liu et al. [19] designed a novel trainable Grid-Dehaze Network (GDN) that indicates how confident the network is about the multi-scale features are learned. Following it, massive studies [6, 16, 36] have pointed out that different features at a range of scales show high importance to infer relative haze thickness according to an individual image. Therefore, the network architecture should exhibit some fashion for efficiently processing and integrating features at various scales.

However, few attempts have been explored to preserve the desired fine image details and strong contextual information. More fundamentally, conventional multi-scale manners [6, 14, 16] do not make full use of image features in the previous layers, making it difficult to transfer local features to other layers, and their performance is easily limited by a bottleneck effect as the multi-scale feature flow will be suppressed [33, 35]. Therefore, investigating the way of capturing information at various scales and constructing more effective hierarchy architecture is essential for image reconstruction. To this end, we perform multi-scale estimation on an end-to-end deep hourglass-structured hierarchical model to remove haze, which flexibly allows efficient information exchange at various scales. The hourglass network has clear advantages over the conventional multi-scale network widely adopted in several vision tasks [5, 22, 30]. For further achieve distinctive features conducive to dehazing, a residual dense attention module is embedded in the backbone network, which can also enhance the robustness of the proposed model. Figure 1 presents one dehazing example on a real-world image, where our method products a satisfactory result against other competing methods in restoring the detailed structure information.

The main contributions of this work are as follows:

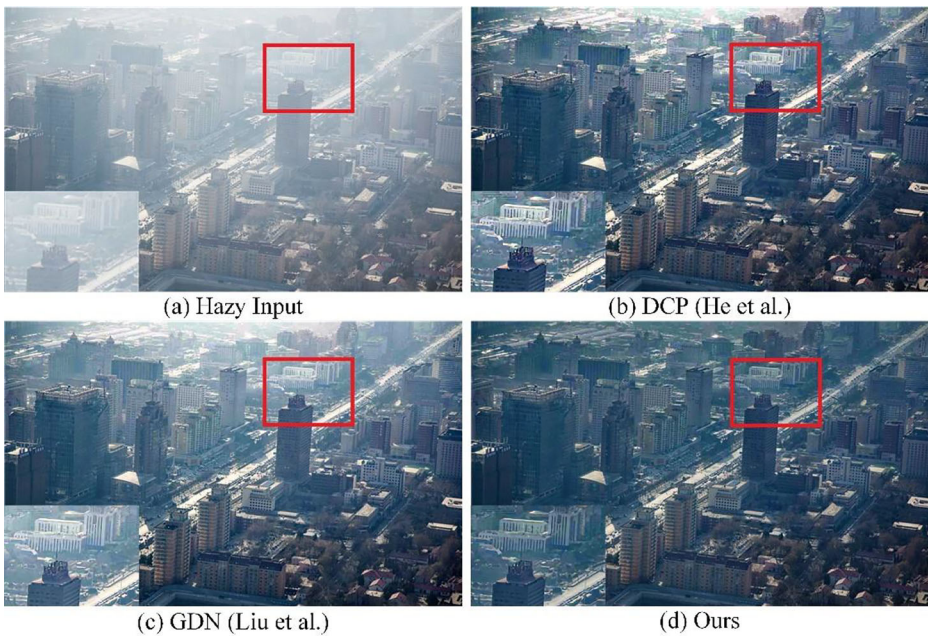


Fig. 1 Dehazing example on a real-world hazy image

- We propose an effective framework to achieve the single image dehazing task, which also provides an effective multi-scale processing solution thanks to its modified hourglass architecture.
- To our knowledge, the Residual Dense Attention Module (RDAM) is first constructed to boost the feature extraction performance to a great extent.
- Extensive experiments are conducted on both synthetic benchmark datasets and real hazy images. The developed method outclasses the existing approaches in visually and quantitatively comparisons.

The remainder of the paper is organized as follows: Section 2 describes an overview of related work. In Section 3, we present the main technical details. Then, Section 4 shows the comprehensive experiments results with discussions. At last, the conclusions are given in Section 5.

2 Related work

In this section, we introduce several proposed image dehazing approaches based on prior-based and learning-based types. In addition, recently developing states of multi-scale feature fusion is presented.

2.1 Single image dehazing

Single image dehazing is an ill-posed issue, and extensive works are carried out for addressing this issue. Based on the model of atmosphere scattering [20, 21], the hazing process is defined:

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

In which x indicates the pixel coordinate, $I(x)$ and $J(x)$ denote the hazy image under observation and the clear image, respectively. There are two critical parameters: $t(x)$ refers to the medium transmission map, and A is termed as the global atmospheric light. Under the homogeneous haze, the map of transmitting process $t(x)$ is written:

$$t(x) = \exp(-\beta d(x)) \quad (2)$$

In which $d(x)$ and β denote the scene depth and the atmosphere scattering coefficient. Since only $I(x)$ is available, the process to recover the haze-free scene $J(x)$ becomes a mathematically ill-posed issue. The existing dehazing approaches are able to fall to prior-based and learning-based strategies, as presented in the following.

2.1.1 Prior-based dehazing

For recovering the haze-free images from hazy images, prior-based approaches largely comply with prior data and assumptions. For instance, the albedo of the scene can receive the estimation to be prior knowledge by referencing [7]. Assuming that the local contrast of hazy image was low, [31] put forward Markov Random Field for maximizing color contrast. For the more reliable calculation of the transmission map, [1] reported the dark channel for quality improvement related to dehaze images. Greater advancements in terms of the dark channel were achieved on [13]. [1, 8] presented their algorithm in a separate manner by complying with the observation that image patches usually show one-dimensional feature distribution in RGB color space. According to conventional prior-based approaches, the assumption was true in limited scenarios, and it was partially restricted.

2.1.2 Learning-based dehazing

Recently, many learning-based approaches are employed to achieve dehazing by exploiting the large-scale datasets and computational power advancement. Early learning-based approaches continuously comply with global atmospheric scattering model and restore haze-free image through the estimation of the global atmospheric light and transmission map. Ren et al. [25] designed a multiscale framework to estimate a transmission map. Cai et al. [2] developed one DehazeNet built with feature layers to predict transmission map. Nevertheless, it exhibits susceptibility to the noise, thereby reducing dehazing performance. For this reason, end-to-end CNNs are developed for directly outputting a clear image based on a hazy input with no use of global atmospheric scattering pattern. Chen et al. [3] employed an encoder-decoder network exhibiting smoothed dilated convolution in the model for gridding artifacts alleviation. Qu et al. [24] converted image dehazing as one image-to-image translating problem, and proposed one optimized pix2pix dehazing framework. However, the mentioned approaches largely comply with several generic framework architectures with no noticeable change, showing inefficiency in terms of image dehazing. The present paper further explores a more efficient framework to promote the utilization and transmission of image feature flows.

2.2 Multi-scale feature fusion

Feature fusion receives extensively applications for designing framework to gain performance through the use of features of a range of layers. Considerable image restoring approaches fuse

features based on dense connecting process [39], feature concatenating process [40] or weighted element-wise summation [3, 41]. In [11], the features at a range of scales undergo the projecting process and concatenating process based on the strided convolutional layer. Though the mentioned approaches can merge multiple features from a range of extents, the concatenation scheme exhibits ineffectiveness to achieve efficient feature extraction. For sharing data of nearby extents, the grid architectures [9, 42] have been developed through the interconnection of the features from nearby extents with convolutional and deconvolutional layers. Liu et al. [19] recently developed one feasible end-to-end trainable grid framework to dehaze images. Inconsistent with [19], one deep hourglass-structured hierarchical model was developed, fusing features under a range of scales significantly. Our work aims to integrate the multi-scale fusion into the hourglass framework to better exploit and mine multi-scale information.

3 Methodology

In this section, we elucidate the developed method, consisting of overall network design, backbone component, and loss function.

3.1 Overall design

The proposed network design largely complies with [22], which initially presented a “stacked hourglass” framework develop for the task of human pose estimation. Figure 2 indicates that the Hourglass is inconsistent with previous designs largely because of its higher symmetrical topological characteristic, enabling the network for capturing and integrating data at various image scales. We continue along this trajectory and further adjust the hourglass architecture to meet the need for dehazing. To our knowledge, this is the first study of developing hourglass-structured dehazing model, which enriches the literature of this field.

Figure 3 gives an overview of our method. The main body, contains three parallel streams, which accordingly feature map number at 3 scales is increased to the double, with values of 16, 32 and 64. We start from a coarse-scale stream as the first level input, and then gradually repeat top-down and bottom-up processing to form more coarse-to-fine scale streams one by one. In each stream, the residual dense attention modules (RDAM) are introduced as the basic component to deeply learn the features for hazy regions reconstruction. Visually, the path of

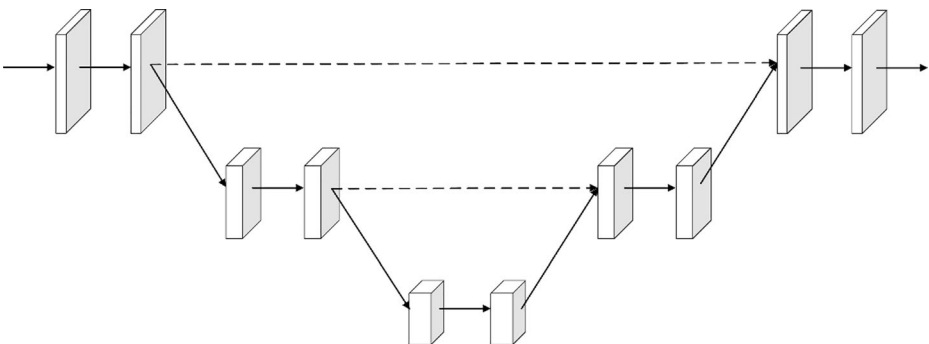


Fig. 2 The illustration of a single hourglass network

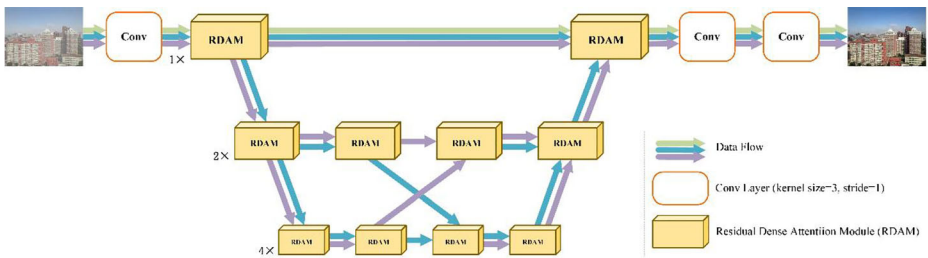


Fig. 3 The overall network framework of the proposed deep hourglass-structured fusion model for dehazing. The network consists of three parallel convolution streams with channel dimensions of 16, 32, 64 at resolution 1, 1/2, 1/4

expansion exhibits symmetry towards the path for contract, yielding an hourglass-shaped frame. This operation enables the model to better reorganize and explore features in width and depth. In all paths, it is worth mentioning that up-sampling and down-sampling are achieved with the use of a range of convolution-based layers for the adjustment of feature map size, instead of conventional bilinear or bicubic interpolation.

Feature fusion in network design is motivated by the need to exploit features from different layers for performance gain. By combining features at different scales, the object and its surroundings can be better represented [12]. To our knowledge, encoder-decoder network information flow or conventional multi-scale network information flow easily face the bottleneck influence resulting from the hierarchical architecture, while the developed method can effectively alleviate this issue via connections at a range of scales via up-sampling/down-sampling process. It is noteworthy that the approach connects coarse-to-fine scale pipelines in parallel rather than in series as done in most existing multi-scale solutions [40]. Throughout the whole process, we perform repetitive multi-scale fusions through the exchange of the data across the paralleling multi-scale pipelines. Since the final output is an aggregation of the input maps, this study finds that contracting path integrating the feature and multi-scale information facilitates statistical features learning to remove haze of images.

3.2 Residual dense attention module

Feature extraction refers to the critical step in image reconstruction. For achieving accurate clear image estimation, RDAM is introduced to deeply learn detailed features, as inspired by [18, 34] with the aims to recover more image details with better color information preserving. Figure 4 indicates that one RDAM consists of N dense residual blocks (RDB) in series with channel-wise attention, where N is set to 3. Since RDB is capable of more fully exploiting both the dense block and the residual block, it is adopted to link all matching feature maps, and reuse low-level features at high levels to enhance dehazing performance, thereby ensuring maximum information flow between the layers of the network. However, these features fused after stacking the RDBs will inevitably be confused with the artifacts generated during CNN, which may reduce the final performance noticeably. The channel attention mechanism [34] clarifies this problem, and RDAM only employs the channel-wise attention once, thereby greatly increasing model efficiency.

To better explain the working mechanism of the block, we provide the decomposed structure of RDB in Fig. 4. Specifically, in each RDB, the first four convolutional layers are adopted to elevate the number of feature maps, while the last convolutional layer is employed

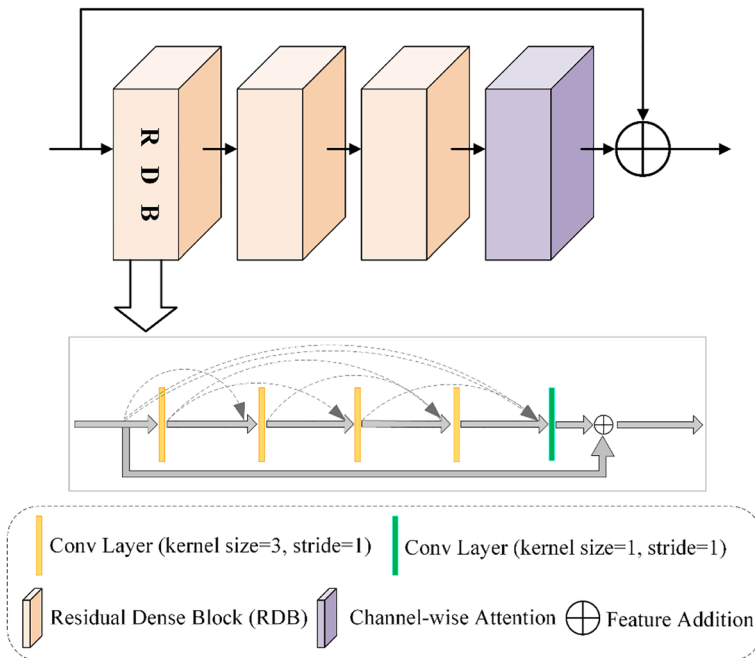


Fig. 4 The architecture of residual dense attention module (RDAM)

to aggregate these feature maps. Subsequently, its output is merged with the input of RDB via channel-wise addition. Note that, the growth rate of RDB is set to 16. Except for the 1×1 convolutional layer in each RDB, all convolutional layers adopt rectified linear unit (ReLU) as the activation function.

Since different channels have totally different weighted information [23], channel attention is added after the stacked RDBs to acquire necessary local features. First, the channel-wise global spatial information is transformed into a channel descriptor based on global average pooling.

$$z_c = H_p(F_c) = \frac{1}{H \times W} \sum_{i=1}^n \sum_{j=1}^n X_c(i, j) \tag{3}$$

where $X_c(i, j)$ denotes the value of c -th channel X_c at the location (i, j) and H_p stands for global pooling.

In order to get the weights of various channels, the Sigmoid and ReLU activation functions are used following with two convolution layers.

$$CA_c = \sigma(\text{Conv}(\delta(\text{Conv}(z_c)))) \tag{4}$$

Eventually, we element-wise multiply the input F_c and the weights information CA_c .

$$F_c^* = F_c \otimes CA_c \tag{5}$$

3.3 Loss function

By optimizing the model parameters, the linear combination of two losses is taken as the optimization objective here, the MSE loss and the perceptual loss. Mathematically,

$$L_{mse} = \|J - J_{GT}\|_2 \quad (6)$$

$$L_{perceptual} = \sum_{l \in \{4,9,16,23\}} \|vgg_l(J) - vgg_l(J_{GT})\|_1 \quad (7)$$

where J and J_{GT} respectively represent the predicted image and the ground truth for supervision; $vgg_l(\cdot)$ denotes the l^{th} features obtained by the layer within the VGG-16 model. Lastly, the total loss function can be written as:

$$L_{total} = L_{mse} + \lambda L_{perceptual} \quad (8)$$

where λ denotes a loss weight.

4 Experimental results

In this section, we have conducted extensive experiments to compare the performance of different dehazing algorithms. Besides, ablation studies are also conducted to analyze the configuration of the proposed model.

4.1 Experimental settings

The available benchmark synthetic datasets, named RESIDE [17] and HazeRD [38], is employed for assessments. RESIDE contains different hazy images in outdoor and indoor scenes, which is generated by atmospheric scattering model. For testing, synthetic Objective Testing Set (SOTS) is the test subset of the RESIDE, containing 1000 hazy images. HazeRD contains 75 synthesized hazy images with real fog conditions. As the ground-truths available, the common PSNR and SSIM [32] act as the metrics of this study, with higher value for better quality. Furthermore, 30 real hazy images are selected from the Internet data to verify the generalization ability of different models.

In the training process, similar to [29], RGB images directly act as input instead of using image patches. The implementation here is developed in PyTorch framework and performed on one NVIDIA Tesla V100 GPU. The final model is trained for 120 epochs. The learning rate is set to 0.001, and then attenuated by half every 20 epochs. Given the loss term, the weight parameter is set as $\lambda = 0.04$, empirically. The Adam optimizer is adopted with a batch size of 16, where β_1 and β_2 respectively take the default values of 0.9 and 0.999.

4.2 Quantitative results

We compare our models with some existing classic and popular approaches, including DCP [10], NLD [1], DehazeNet [2], MSCNN [25], AOD-Net [14], GFN [26], GCA-Net [3] and GDN [19]. Table 1 lists the average values between each pair of dehazed result and haze-free image. From the table, we can observe that our Hourglass-DehazeNet achieves the best performance on the SOTS [18] and HazeRD [34] datasets, which indicates that the best detailed recovery and the least noise impact are obtained by our hourglass-structured fusion model. Compared with SOTS dataset, HazeRD dataset is more challenging [27], and what's

Table 1 Quantitative comparison results on the SOTS [18] and HazeRD [34] dataset

Method	SOTS		HazeRD	
	PSNR	SSIM	PSNR	SSIM
DCP [10]	18.75	0.8590	15.58	0.7673
NLD [1]	17.28	0.7503	16.20	0.7278
MSCNN [2]	17.57	0.8115	15.87	0.7949
DehazeNet [25]	21.14	0.8500	15.56	0.7552
AOD-Net [14]	18.80	0.8342	16.85	0.8331
GFN [26]	24.11	0.8996	16.54	0.8024
GCA-Net [3]	28.13	0.9455	16.62	0.8195
GDN [19]	31.51	0.9804	17.33	0.8384
Ours	33.54	0.9849	19.08	0.8719



Fig. 5 Qualitative comparison results on the SOTS [18] dataset

exciting is that the SSIM and PSNR produced by our approach are higher than GDN by up to 0.03 and 1.75 dB, respectively.

4.3 Qualitative results

Besides quantitative comparisons, several visual examples are taken as comparison, as illustrated in Figs. 5 and 6. Here, it is noteworthy that the visualization results of NLD, MSCNN and DehazeNet are not presented for their relatively inferior performance and the space limitation. By observing zoomed parts of image, DCP is clearly indicated to make the background darker and over saturated. We can figure out that the AOD-Net and GFN underestimate the haze level and leave lots of haze residuals. GCA-Net can remove the haze well but bring about color artifacts and amplified noise. Moreover, blurry outcomes achieved by GDN with lost details reveal the defect of the current



Fig. 6 Qualitative comparison results on the HazeRD [34] dataset

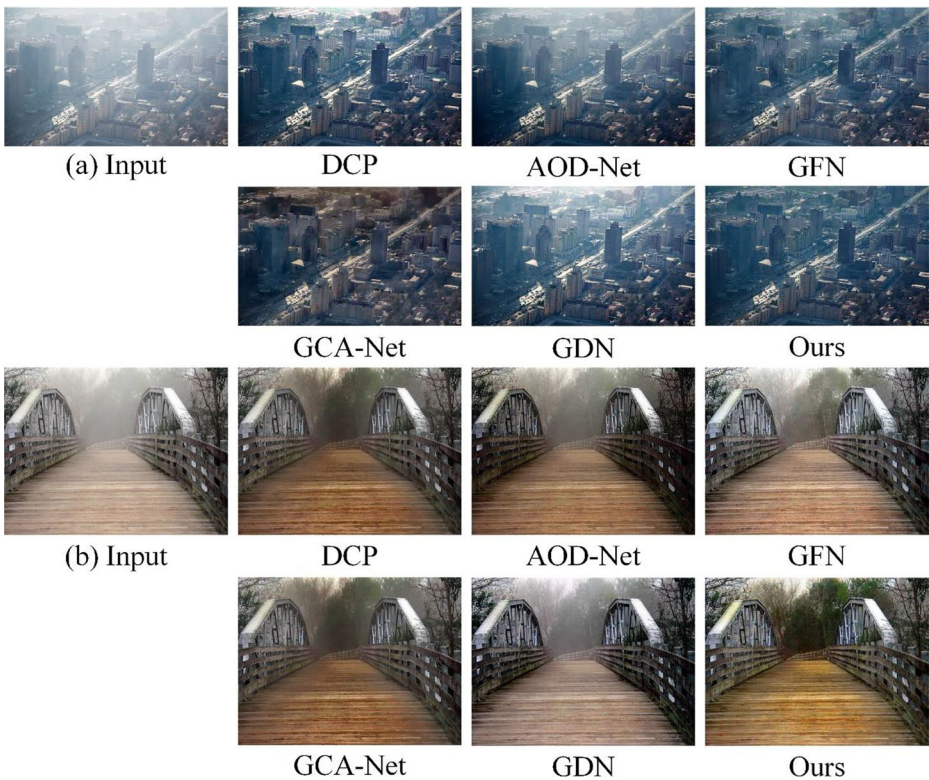


Fig. 7 Qualitative comparison results on the real-world hazy images

approach. In contrast, the proposed method can restore images with better color fidelity and sharper details, which are confirmed to be closer to the ground-truths. Accordingly, the simple but effective idea of this study for such a modified hourglass architecture is clarified effectively, and the designed basic model is confirmed.

Moreover, the generalization ability of the proposed algorithm on real images is evaluated. From Fig. 7, we can see that the amount of haze reduced by the existing techniques is limited as compared to that of our method. In contrast, our proposed approach can remove almost all the haze in diverse haze distribution with complex background. In addition, another benefit can be found is being good at restoring the detailed structure information. For the overall subjective perception, our model exhibits impressive restoration performance.

4.4 Ablation studies

The effectiveness of configuration and parameters in the network proposed is verified through ablation studies, which is based on the SOTS dataset and conducted in the same environment.

Table 2 Ablation study on the number of residual dense blocks in RDAM

Block No.	N=1	N=2	N=3	N=4
PSNR	31.27	32.60	33.54	33.56
SSIM	0.9786	0.9812	0.9849	0.9848

Table 3 Ablation study on channel-wise attention

	Baseline (w/o channel attention)	Our method
PSNR	32.86	33.54
SSIM	0.9827	0.9849

Number of residual dense blocks in RDAM The experiments are performed under different numbers of residual dense blocks (RDB) to explore their influences. To be more specific, the number of RDB is set to $N \in \{1, 2, 3, 4\}$. Table 2 reports the performances. PSNR and SSIM values grow as blocks increase; however, when N reaches 3 and extra time is cost, these values undergo limited increment. Therefore, to balance effectiveness and efficiency, the default number of N is set to 3.

Effectiveness of channel-wise attention Afterward, different variants of RDB are studied to explore the effectiveness of the introduced attention mechanism. The channel-wise attention is removed to build the baseline module. Table 3 demonstrates the ability of the feature attention module to improve PSNR and SSIM. When the channel-wise attention is adopted, the performance is the best and the total gain over the baseline is 0.68 dB in term of PSNR. It indicates its role in promoting haze removal.

5 Conclusion

In this work, an effective deep hourglass-structured fusion model is developed for processing the single image dehazing task. Our approach achieves the abundant representation of features at various scales, where the abundant multi-scale haze-relevant features captured from the residual dense attention module is gradually fused along the parallel layers and stages of the modified hourglass architecture. Furthermore, a channel-wise attention mechanism is presented for improving network performance, which treats channel unequally, thereby showing extra flexibility in feature fusion. As proved by extensively conducted studies, the proposed method outclasses several existing approaches on the synthetic benchmark dataset and the practical ones. In future, we plan to employ the developed network idea into other image restoration tasks as impacted by the generic characteristic exhibited by its building parts.

References

1. Berman D, Treibitz T, Avidan S (2016) Non-local image Dehazing. In: 2016 IEEE/CVF conference on computer vision and pattern recognition. pp 1674–1682. <https://doi.org/10.1109/cvpr.2016.185>
2. Cai B, Xu X, Jia K, Qing C, Tao D (2016) DehazeNet: An end-to-end system for single image haze removal. *IEEE Trans Image Process* 25(11):5187–5198. <https://doi.org/10.1109/TIP.2016.2598681>
3. Chen D, He M, Fan Q, Liao J, Zhang L, Hou D, Yuan L, Hua G (2019) Gated context aggregation network for image Dehazing and Deraining. In: 2019 IEEE winter conference on applications of computer vision, pp 1375–1383. <https://doi.org/10.1109/wacv.2019.00151>
4. Chen X, Li YF, Dai LG, Kong CH (2021) Hybrid high-resolution learning for single remote sensing satellite image Dehazing. *IEEE Geosci Remote Sens Lett* 19:1–5. <https://doi.org/10.1109/LGRS.2021.3072917>

5. Chen X, Huang YF, Xu L (2021) Multi-scale hourglass hierarchical fusion network for single image Deraining. In: 2021 IEEE/CVF conference on computer vision and pattern recognition workshops, pp 872–879
6. Dong H, Pan J, Xiang L, Hu Z, Zhang X, Wang F, Yang M-H (2020) Multi-scale boosted dehazing network with dense feature fusion. In: 2020 IEEE/CVF conference on computer vision and pattern recognition, pp 2154–2164. <https://doi.org/10.1109/CVPR42600.2020.00223>
7. Fattal R (2008) Single image dehazing. *ACM Trans Graph* 27(3):1–9. <https://doi.org/10.1145/1360612.1360671>
8. Fattal R (2014) Dehazing using color-lines. *ACM Trans Graph* 34(1):1–14. <https://doi.org/10.1145/2651362>
9. Fournre D, Emonet R, Fromont E, Muselet D, Tremeau A, Wolf C (2017) Residual conv-deconv grid network for semantic segmentation arxiv: 1707.07958
10. He K, Sun J, Tang X (2011) Single image haze removal using dark channel prior. *IEEE Trans Pattern Anal Mach Intell* 33(12):2341–2353. <https://doi.org/10.1109/TPAMI.2010.168>
11. Huang G, Chen D, Li T, Wu F, Van Der Maaten L, Weinberger K (2018) Multi-scale dense networks for resource efficient image classification arxiv: 1703.09844
12. Jiang K, Wang Z, Yi P, Chen C, Huang B, Luo Y, Ma J, Jiang J (2020) Multi-scale progressive fusion network for single image Deraining. In: 2020 IEEE/CVF conference on computer vision and pattern recognition, pp 8343–8352. <https://doi.org/10.1109/CVPR42600.2020.00837>
13. Levin A, Lischinski D, Weiss Y (2008) A closed-form solution to natural image matting. *IEEE Trans Pattern Anal Mach Intell* 30(2):228–242. <https://doi.org/10.1109/TPAMI.2007.1177>
14. Li B, Peng X, Wang Z, Xu J, Feng D (2017) AOD-Net: All-in-one dehazing network. In: 2017 IEEE international conference on computer vision (ICCV), Venice, Italy, pp. 4780–4788. <https://doi.org/10.1109/ICCV.2017.511>
15. Li RD, Pan JS, Li ZC, Tang JH (2018) Single image Dehazing via conditional generative adversarial network. In: IEEE/CVF conference on computer vision and pattern recognition, New York, pp. 8202–8211. <https://doi.org/10.1109/cvpr.2018.00856>
16. Li Y, Miao Q, Liu R, Song J, Quan Y, Huang Y (2018) A multi-scale fusion scheme based on haze-relevant features for single image dehazing. *Neurocomputing* 283:73–86. <https://doi.org/10.1016/j.neucom.2017.12.046>
17. Li B, Ren W, Fu D, Tao D, Feng D, Zeng W, Wang Z (2019) Benchmarking single image dehazing and beyond. *IEEE Trans Image Process* 28(1):492–505. <https://doi.org/10.1109/TIP.2018.2867951>
18. Lim B, Son S, Kim H, Nah S, Lee KM (2017) Enhanced deep residual networks for single image super-resolution. In: 2017 IEEE/CVF conference on computer vision and pattern recognition workshops, pp 1132–1140. <https://doi.org/10.1109/cvprw.2017.151>
19. Liu X, Ma Y, Shi Z, Chen J (2019) GridDehazeNet: attention-based multi-scale network for image Dehazing. In: 2019 IEEE/CVF international conference on computer vision, pp 7313–7322. <https://doi.org/10.1109/iccv.2019.00741>
20. Narasimhan SG, Nayar SK (2002) Vision and the atmosphere. *Int J Comput Vis* 48(3):233–254. <https://doi.org/10.1023/a:1016328200723>
21. Nayar SK, Narasimhan SG (1999) Vision in bad weather. *The Proceedings of the Seventh IEEE International Conference on Computer Vision* 2:820–827
22. Newell A, Yang K, Deng J (2016) Stacked hourglass networks for human pose estimation. In: proceedings of the European conference on computer vision (ECCV), pp 483–499. https://doi.org/10.1007/978-3-319-46484-8_29
23. Qin X, Wang Z, Bai Y, Xie X, Jia H (2020) FFA-net: feature fusion attention network for single image dehazing. In: AAAI, pp. 11908–11915
24. Qu Y, Chen Y, Huang J, Xie Y, Soc IC (2019) Enhanced Pix2pix Dehazing network. In: 2019 IEEE/CVF conference on computer vision and pattern recognition, pp 8152–8160. <https://doi.org/10.1109/cvpr.2019.00835>
25. Ren W, Liu S, Zhang H, Pan J, Cao X, Yang MH (2016) Single image dehazing via multi-scale convolutional neural networks. *European Conference Comput Vis* 9906:154–169. <https://doi.org/10.1007/978-3-319-46475-6-10>
26. Ren W, Ma L, Zhang J, Pan J, Cao X, Liu W, Yang MH (2018) Gated fusion network for single image dehazing. In: 2018 IEEE/CVF conference on computer vision and pattern recognition, Salt Lake City, UT, USA, pp. 1–9. <https://doi.org/10.1109/cvpr.2018.00343>
27. Shao Y, Li L, Ren W, Gao C, Sang N (2020) Domain adaptation for image dehazing. In: 2020 IEEE/CVF conference on computer vision and pattern recognition, pp 2805–2814. <https://doi.org/10.1109/cvpr42600.2020.00288>
28. Sharma T, Agrawal I, Verma NK (2020) CSIDNet: compact single image dehazing network for outdoor scene enhancement. *Multimed Tools Appl* 79(41–42):30769–30784. <https://doi.org/10.1007/s11042-020-09496-z>

29. Shen J, Li Z, Yu L, Xia G-S, Yang W (2020) Implicit euler ODE networks for single-image dehazing. In: 2020 IEEE/CVF conference on computer vision and pattern recognition workshops, pp 877–886. <https://doi.org/10.1109/cvprw50498.2020.00117>
30. Sun K, Xiao B, Liu D, Wang J, Soc IC (2019) Deep high-resolution representation learning for human pose estimation. In: 2019 IEEE/CVF conference on computer vision and pattern recognition. pp 5686–5696. <https://doi.org/10.1109/cvpr.2019.00584>
31. Tan RT (2008) Visibility in bad weather from a single image. In: 2008 IEEE/CVF conference on computer vision and pattern recognition, pp 2347–2354. <https://doi.org/10.1109/cvpr.2008.4587643>
32. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process* 13(4):600–612. <https://doi.org/10.1109/tip.2003.819861>
33. Wang C, Zhang M, Su Z, Yao G (2020) Densely connected multi-scale de-raining net. *Multimed Tools Appl* 79(27):19595–19614. <https://doi.org/10.1007/s11042-020-08855-0>
34. Woo S, Park J, Lee J-Y, Kweon IS (2018) CBAM: convolutional block attention module. In: Proceedings of the European conference on computer vision (ECCV), pp 3–19. https://doi.org/10.1007/978-3-030-01234-2_1
35. Yang Y, Zhang D, Huang S, Wu J (2019) Multilevel and multiscale network for single-image super-resolution. *IEEE Signal Process Lett* 26(12):1877–1881. <https://doi.org/10.1109/lsp.2019.2952047>
36. Yeh C-H, Huang C-H, Kang L-W (2020) Multi-scale deep residual learning-based single image haze removal via image decomposition. *IEEE Trans Image Process* 29:3153–3167. <https://doi.org/10.1109/tip.2019.2957929>
37. Zhang H, Patel VM (2018) Densely Connected Pyramid Dehazing Network. In: 2018 IEEE/CVF conference on computer vision and pattern recognition, New York, pp 3194–3203. <https://doi.org/10.1109/cvpr.2018.00337>
38. Zhang Y, Ding L, Sharma G (2017) HAZERD: AN OUTDOOR SCENE DATASET AND BENCHMARK FOR SINGLE IMAGE DEHAZING. In: 2017 24th IEEE international conference on image processing (ICIP). pp 3205–3209
39. Zhang H, Sindagi V, Patel VM (2018) Multi-scale single image Dehazing using perceptual pyramid deep network. In: proceedings 2018 IEEE/CVF conference on computer vision and pattern recognition workshops, pp 1015–1024. <https://doi.org/10.1109/cvprw.2018.00135>
40. Zhang Y, Tian Y, Kong Y, Zhong B, Fu Y (2018) Residual dense network for image super-resolution. In: 2018 IEEE/CVF conference on computer vision and pattern recognition, pp 2472–2481. <https://doi.org/10.1109/cvpr.2018.00262>
41. Zhang X, Dong H, Hu Z, Lai W-S, Wang F, Yang M-H (2019) Gated fusion network for joint image deblurring and super-resolution arxiv: 1807.10806
42. Zung J, Tartavull I, Lee K, Seung HS (2017) An Error Detection and Correction Framework for Connectomics arxiv: 1708.02599

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Yufeng Li¹ · Xiang Chen¹ · Caihua Kong¹ · Longgang Dai¹ · Yufeng Huang¹

Yufeng Li
liyufeng@sau.edu.cn

Caihua Kong
kongcaihua@stu.sau.edu.cn

Longgang Dai
dailonggang@stu.sau.edu.cn

Yufeng Huang
yufengh_sau@sina.com

¹ College of Electronic Information Engineering, Shenyang Aerospace University, Shenyang 110135, China