Multi-priors Guided Dehazing Network Based on Knowledge Distillation

Nian Wang, Zhigao Cui$^{(\boxtimes)}$, Aihua Li, Yanzhao Su, and Yunwei Lan

Xi'an Research Institute of High-tech, Xi'an 710025, China
cuizg10@126.com

Abstract. Single image dehazing is a key prerequisite of high-level computer vision tasks since degraded images seriously affect the recognition ability of computers. Traditional prior-based methods conduct favorable dehazing effect but tend to cause artifacts and color distortions due to inaccurate parameter estimations. By contrast, recent learning-based methods can provide better color fidelity via the supervised training of synthetic images. But unfortunately, these methods always acquire under-dehazed results due to the domain differences between synthetic hazy images and their real-world ones. To combine the merits of these two categories, we propose a multi-priors guided dehazing network (MGDNet) based on knowledge distillation. Specifically, we adopt the dehazed images of dark channel prior and non-local dehazing prior as fake ground truths, and use them to pretrain two teacher networks. Then we build a student network based on encoder-decoder structure, and set up both feature-level and pixel-level knowledge distillation losses to guide the training process of the student network. Experimental results on some real-world datasets widely used in recent works demonstrate that our MGDNet can generate visually appealing images with more discriminative textures and vivid color when compared with the state-of-the-arts.

Keywords: Single image dehazing · Knowledge distillation · Multi-priors guiding

1 Introduction

Particles in the atmosphere absorb and scatter the reflected lights of object and result in poor image visibility, which hinders the performance of high-level computer vision tasks [1]. Hence, as a key prerequisite, single image dehazing has been widely studied in the latest decade, which can be roughly divided into model-based methods and model-free methods [2].

Traditional model-based methods estimate the unknown atmospheric light and transmission maps by the statistical rules of haze-free images, which include dark channel prior (DCP) [3], color-lines prior (CLP) [4], color attenuation prior (CAP) [5], and non-local dehazing (NLD) [6]. These methods achieve favorable dehazing effect and generalization ability, but tend to cause some color distortion and artifacts since unilateral hypothesis cannot maintain the accuracy of parameter estimations in various scenes.

To this end, recent model-based methods utilize convolutional neural networks (CNNs) to estimate the atmospheric light and transmission maps respectively [7, 8] or simultaneously [9, 10]. These learning-based methods estimate parameters by data driving rather than man-made priors, and thus acquire more visually pleasing images. However, the atmospheric scattering model is just an ideal model, which influences the convergence of networks and restricts the final dehazing performance [11].

More recently, learning-based methods [12–16] tend to avoid the atmospheric scattering model and adopt an end-to-end training strategy (directly building the mapping between hazy images and haze-free images) to acquire high quality results. However, due to the huge differences between the features of hazy images and their haze-free ones, model-free methods always expand the network depths and scales to enhance feature extraction ability, which results in large computational consumption. Moreover, these methods fail to dehaze when applied to real scenes, mainly because networks trained on synthetic dataset cannot fit in uneven haze distribution and complex illumination existing in real scenes. To this end, some works [17–19] start to combine prior-based methods and model-free methods to reduce the differences between synthetic domain and real domain, which achieve better dehazing effect when applied to real-world images.

In this paper, we propose a multi-priors guided dehazing network (MGDNet) based on knowledge distillation. Different from a recent work [20], we pretrain two teacher networks by minimizing the losses between hazy images and supervised images (dehazed images of dark channel prior and non-local dehazing), and then teach a student network to learn their features by minimizing both feature-level and pixel-level distillation loss. Considering that the supervisions of teacher networks contain some color distortion, we utilize discrete wavelet transform (DWT) to get the high-frequency and low-frequency of the outputs of teacher networks and only use the high-frequency part to build the pixel-level distillation loss.

Comparative experiments on some real-world hazy images show that our MGDNet performs favorably against the state-of-the-arts, which validates that guiding with the partially correct teacher networks (the supervisions are dehazed images of DCP and NLD rather than ground truths) can effectively improve the dehazing ability in real scenes. In addition, these added negative information from teacher networks can be refined by the training process and the student network finally acquire dehazed images with more vivid color.

2 Related Work

2.1 Model-Based Methods

Model-based methods estimate the atmospheric light and transmission maps, and then restore dehazed images by atmospheric scattering model. Early model-based methods, also called prior-based methods, adopt statistical assumptions concluded from haze-free images to estimate the atmospheric light and transmission maps. For example, dark channel prior (DCP) [3] assumes clear RGB images have low intensity in at least one channel, and quickly acquires these two parameters based on the theory. Color-lines prior (CLP) [4] constructs a local formation model to recover the transmission map based on the lines offset. In addition, Color attenuation prior (CAP) [5] builds a linear relationship

among color, haze concentration and scene depth to estimate the atmospheric light and transmission maps. Differently, another method NLD [6] estimates the transmission map by hundreds of distinct colors. Above prior-based methods dehaze favorably and have strong generalization in real scenes but tend to cause artifacts, halos and color distortion since unilateral assumption cannot estimate accurate atmospheric light and transmission maps in various scenes. To this end, recent model-based methods tend to estimate the atmospheric light and transmission maps by convolutional neural networks (CNNs). For example, some works estimate transmission maps by stacked CNN [7] or multiscale CNN [8]. Moreover, to avoid the cumulative error of two times estimation, AOD-Net [9] sets a linear equation to combine the atmospheric light and transmission map into a parameter $K(x)$. Another method DCPDN [10] embeds the atmospheric scattering model into CNN, which directly acquires dehazed images by the joint estimation of atmospheric light and transmission maps. However, the atmospheric scattering model, as a simplified mathematical model, cannot completely replace the formation process of haze. Hence, model-based methods cannot acquire high quality results and still suffer from some color and illumination changes.

2.2 Model-Free Methods

Model-free methods, also called end-to-end methods, directly establish the mapping between hazy images and clear images instead of using atmospheric scattering model. Due to the huge gap between the features of hazy images and clear images, model-free methods often increase network depths and scales to enhance the feature extraction ability. For example, FFA [12] and DuRN [13] build a deep network based on residual blocks, and directly recover dehazed images by merging features from convolutional layers in different depths. GFN [14] utilizes white balance (WB), contrast enhancing (CE), and gamma correction (GC) to derive three subimages of the hazy input, and directly recovers dehazed images by using learned confidence maps to fuse these three subimages. Moreover, EPDN [15] acquires high contrast results by the adversarial training between a multiscale generator and discriminator. MSBDN [16] adopts back-projection feedback to connect non-adjacent layers, which reduces the loss of spatial information during sampling and improves the resolution of restored results. However, due to lacking of the knowledge to real-world haze, above networks conduct poor dehazing performance in real scenes. To this end, DANet [17] builds a bidirectional translation network to solve the domain adaptation problem, and acquires visually pleasing results on both synthetic and real scenes. RefineDNet [18] embeds DCP in CNN-based method, and adopts an adversarial training between unpaired real-world images to improve dehazing effect in both synthetic and real scenes. PSD [19] uses multiple prior losses to guide the training process, which acquires high contrast results in real scenes but tend to overenhance images. Differently, KDDN [20] pretrains a reconstruction network of clear images, and adopts the intermediate features to guide the training process of a dehazing network.

3 Proposed Method

As shown in Fig. 1, considering that dark channel prior (DCP) and non-local dehazing (NLD) dehaze favorably in real scenes, we dehaze images by these two prior-based

methods and use them as fake ground truths to pretrain two teacher networks. During the training process of student network, we use the features of teacher networks to guide the student network, and make it achieve favorable dehazing effect in real scenes.
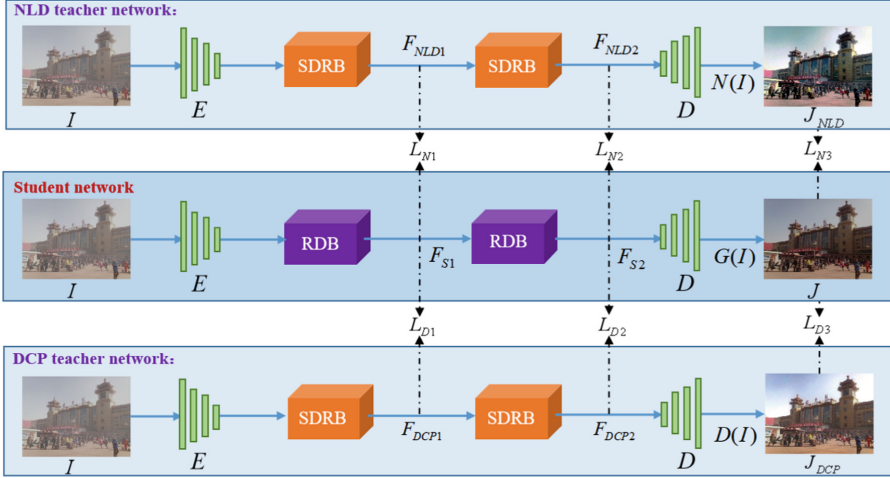

Fig. 1. The architecture of the proposed MGDNet

3.1   Teacher Networks

The DCP and NLD teacher networks have the same architecture, which are based on classic encoder-decoder architecture. As shown in Fig. 1, we first extract features on four scales by an encoder E containing four convolutions. The first convolution preliminarily extracts features of hazy inputs and change the shape from $256 \times 256 \times 3$ to $256 \times 256 \times 64$, and the following three convolutions sequentially adjust the shape of features to $128 \times 128 \times 128$, $64 \times 64 \times 256$ and $32 \times 32 \times 512$, respectively. Moreover, considering paper [21] has shown that applying dilated convolutions into bottleneck layers of encoder-decoder structure can effectively alleviate the generation of artifacts. Hence, we design smoothed dilated residual block (SDRB) and add two SDRBs in the bottleneck layers of these two teacher networks. After that, a decoder D, consists of four deconvolutions, unsamples features to the shape of the corresponding layer in the encoder E, and finally outputs the dehazed images.

As shown in Fig. 2(a), the SDRB consists of two smoothed dilated convolutions (SDC) [22] and a residual connection [23], and each SDC contains a ShareSepConv, a 3×3 convolution and a ReLU function. The ShareSepConv (Separable and Shared Convolution) performs as a preprocessing module, which builds a connection between non-adjacent regions and solves the spatial discontinuity caused by the expansion of receptive field. The theories and details of ShareSepConv can be seen in paper [22]. The 3×3 convolution sets the dilation as 2 to expand the receptive fields, and enhances the perception ability to global features. Finally, the ReLU function improves network nonlinearity and the residual connection after two SDCs enhances feature flow.
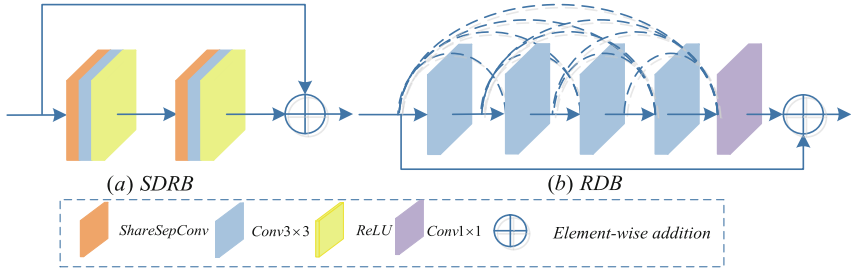
Fig. 2. The structure of SDRB and RDB.

3.2 Student Networks

The student network is a dehazing network trained by synthetic hazy images, which has the similar structures to the teacher networks. As shown in Fig. 1, the student network is still based on a encoder-decoder structure but two residual dense blocks (RDBs) are applied to the bottleneck layers. RDB [24] combines the advantages of residual connection and densely connected network [25], which extracts structures effectively and helps the feature backpropagation. As shown in Fig. 2 (b), these two RDBs contain four 3×3 convolutions and one 1×1 convolution. All 3×3 convolutions are densely connected to avoid the loss of structure information extracted by shallower layers, and then the 1×1 convolution merges these abundant features to provide clear texture perception.

3.3 Overall Loss Function

Recent research [26] has shown that the combination of pixel-wise loss and feature-wise loss can effectively accelerate network training. Hence, for the training of MGDNet, the overall loss function contains L1 loss, perceptual loss, and distillation loss, which can be expressed as Eq. (1):

$$L_{loss} = L_1 + L_{per} + \lambda L_{diss} \tag{1}$$

where L_1, L_{per} and L_{diss} denotes the L1 loss, perceptual loss and distillation loss, respectively. λ is a trade-off coefficient to balance the effect of learning-based method and prior-based method. As shown in Fig. 3, our method can effectively improve the dehazing effect while maintain the color fidelity when setting λ as 1.
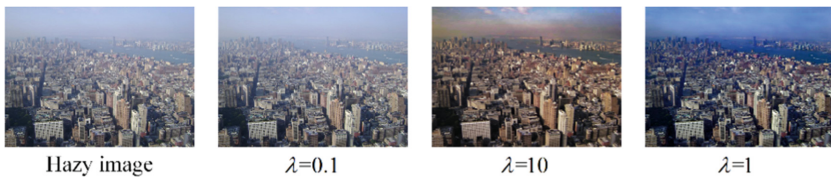

Fig. 3. The results when setting different λ.

L1 Loss. L1 loss (mean absolute error) can rapidly minimize the feature differences between hazy images and clear images by per-pixel comparison, thus we add L1 loss for network training. Different from L2 loss (mean squared error), L1 loss trains networks more stably, which can be expressed as Eq. (2):

$$L_1 = \|J - G(I)\|_1 \tag{2}$$

where J represents haze-free images and $G(I)$ represents the dehazed images of student network.

Perceptual Loss. Perceptual loss [27] compares two images by perceptual and semantic differences, which effectively helps the network restore more vivid images. In this paper, we pretrain VGG19 network on the ImageNet and extract the features of the convolutions in number 2, 7, 15, 21, and 28 (the last convolution of each scale) to calculate loss, which can be expressed as Eq. (3):

$$L_{per} = \sum_{i=1}^{5} \frac{1}{C_i H_i W_i} \|\Phi_i(J) - \Phi_i(G(I))\|_1 \tag{3}$$

where J presents clear images and $G(I)$ represents dehazed images generated by student network. $\Phi_i(J)$ and $\Phi_i(G(I))$ respectively represent the five scales perceptual features of the dehazed images and clear images extracted from the trained VGG19 network. C_i, H_i, and W_i represent the number of channels, height, and width of feature maps.

Distillation Loss. As shown in Fig. 1, to make the trained student network conduct strong generalization ability to real scenes, we pretrain two teacher networks by minimize the L1 loss between $N(I)$ ($D(I)$) and the dehazed images of NLD (DCP) (named as J_{NLD} (J_{DCP})), respectively. Then we adopt the pretrained networks to optimize the training of student network by both of feature-level losses (L_{N1}, L_{D1}, L_{N2}, L_{D2}) and pixel-level distillation losses (L_{N3}, L_{D3}). For feature-level guiding, we output the features after each SDRB (RDB) in the teacher networks (student network) by an extra 3×3 convolution. For pixel-level guiding, considering that the supervisions (dehazed images of NLD and DCP) contain some negative information such as color and illumination distortion, we adopt discrete wavelet transform (DWT) in paper [28] to distinct the high-frequency and low-frequency parts of the outputs of teacher networks, and only the high-frequency images are sent to guide the training process of student network. Hence, the whole distillation loss can be expressed as Eq. (4):

$$\begin{aligned} L_{diss} = \|F_{NLD1} - F_{S1}\|_1 + \|F_{DCP1} - F_{S1}\|_1 + \|F_{NLD2} - F_{S2}\|_1 + \|F_{DCP2} - F_{S2}\|_1 \\ + \|DWT_h(N(I)) - DWT_h(G(I))\|_1 + \|DWT_h(D(I)) - DWT_h(G(I))\|_1 \end{aligned} \tag{4}$$

where F_{DCP1}, F_{DCP2} and F_{NLD1}, F_{NLD2} denote the features extracted from each SDRB of DCP teacher network and NLD teacher network, respectively. F_{S1} and F_{S2} denote the features extracted from each RDB of student network. $DWT_h(\cdot)$ is a high pass DWT, which extracts the structures and textures of input images for pixel-level guiding and makes the student network $G(\cdot)$ avoid the negative low-frequency information (color and illumination distortion) from dehazed images $N(I)$ and $D(I)$.

4   Experiments

In this section, we conduct some experiments on real-world images to show that the proposed MGDNet performs better than some state-of-the-arts. These methods include DCP [3], NLD [6], DANet [17], RefineDNet [18], PSD [19] and KDDN [20]. All these methods are learning-based methods except DCP and NLD, which are two prior-based methods. Moreover, DANet, RefineDNet and PSD are prior-combined methods, and KDDN is a dehazing network using knowledge distillation similar to our MGDNet.

4.1   Dataset

To effectively train our MGDNet, we adopt the Indoor Training Set (ITS) in Realistic Single Image Dehazing (RESIDE) [29], which is a synthetic indoor training set containing 13990 hazy images and the corresponding haze-free images. We test the MGDNet and all comparative methods on IHAZE [30] and OHAZE [31] to verify the dehazing performance, which contains 5 indoor and outdoor paired testing images, respectively. Moreover, some real-world images in [17] and [32] are also adopted to further verify the generalization of MGDNet in real scenes.

4.2   Implementation Details

The proposed MGDNet is trained and tested in the PyTorch framework. During the training, we randomly crop local regions (256×256) of input paired images, and randomly flip or rotate them to enhance the diversity of training dataset. The training batch size is set to 4, and we train the MGDNet 30 epochs. To accelerate the training process, we use the Adam optimizer [33] and adopt a default value for the attenuation coefficient β_1 and β_2 being 0.9 and 0.999 respectively. Moreover, we set the initial learning rate to 0.0002, and decrease it to half every five epochs. All the experiments are implemented on a PC with two RTX 2080Ti GPUs.

4.3   Comparisons with State-of-the-Art Methods

Results on IHAZE and OHAZE. The comparison results are shown in Table 1, where the values are the average PSNR and SSIM of five indoor and outdoor testing images of IHAZE [30] and OHAZE [31], respectively. For IHAZE, the proposed MGDNet achieves the second-best performance of image dehazing in terms of both PSNR and SSIM. For OHAZE, the proposed MGDNet also achieves the second-best PSNR, and enhance SSIM by 0.02 when compared with second-best method DANet. These data show the DANet and our MGDNet conduct better dehazing effect on these two datasets. Also, we notice that prior-based methods DCP and NLD perform poorly in terms of both PSNR and SSIM, which shows the generated artifacts and color changes seriously reduce the quality of the dehazed images.

Table 1. Comparison of the state-of-the-art dehazing methods on IHAZE and OHAZE. Number in red and bule represent the best and second-best results, respectively.

Method		DCP	NLD	DANet	KDDN	PSD	RefineDNet	Ours
IHAZE	PSNR	12.49	13.57	16.23	13.42	13.67	15.9	16.02
	SSIM	0.58	0.59	0.72	0.62	0.63	0.75	0.73
OHAZE	PSNR	14.95	15.24	18.32	16.28	15.72	17.26	18.07
	SSIM	0.67	0.69	0.74	0.71	0.68	0.71	0.76

Results on Natural Hazy Images. Considering that hazy images acquired by haze machine may not completely verify the dehazing ability in real scenes, we further test all these methods on some real-world images in [17] and [33]. As shown in Fig. 4, prior-based methods DCP and NLD dehaze favorably but tent to cause color distortion and artifacts, which shows these methods have excellent generalization ability in real scenes. By contrast, learning-based methods tend to acquire under-dehazed results due to lacking of the knowledge to real-world hazy images. KDDN fails to dehaze in these scenes and a large amount of residual haze degrades the visibility of the generated results. More importantly, DANet cannot remove haze thoroughly and leads to obvious color changes although it shows favorable performance in IHAZE and OHAZE, which verifies this method cannot fit in natural scenes. The results of PSD suffer from severe color and illumination distortion, and some local regions have residual haze. Moreover, RefineDNet dehazes effectively in these scenes and acquires visually pleasing results. However, the results still contain some residual haze in local regions especially when applied to images with colorful textures. Better than above methods, the proposed MGDNet recovers high quality results with discriminative structures and vivid color, which shows it has strong dehazing effect in real scenes by the guiding of DCP and NLD methods. Moreover, compared to the dehazed images of DCP and NLD, we also notice that the results of MGDNet alleviate color distortion and artifacts by the training of synthetic images.

4.4   Ablation Study

To demonstrate the effectiveness of each module, we conduct an ablation study by the combination of four factors: DCP teacher network (DCP), NLD teacher network (NLD), pixel-wise distillation loss (PDL), discrete wavelet transform (DWT). We construct the following variants with different component combinations: (1) Student: only the student network is used; (2) Student+DCP: only the DCP teacher network is used to guide the student network by feature-wise distillation loss. (3) Student + NLD: only.

the NLD teacher network is used to guide the student network by feature-wise distillation losses. (4) Student+DCP+NLD: both the DCP and NLD teacher networks are used to guide the student network by feature-wise distillation losses. (5) Student+DCP+NLD+PDL: two teacher networks guide the student network by both feature-wise distillation losses and pixel-wise distillation losses (6) Student+DCP+NLD+PDL+DWT (Ours): a DWT module is applied before pixel-wise guiding, and only the high-frequency parts of DCP and NLD teacher networks are used
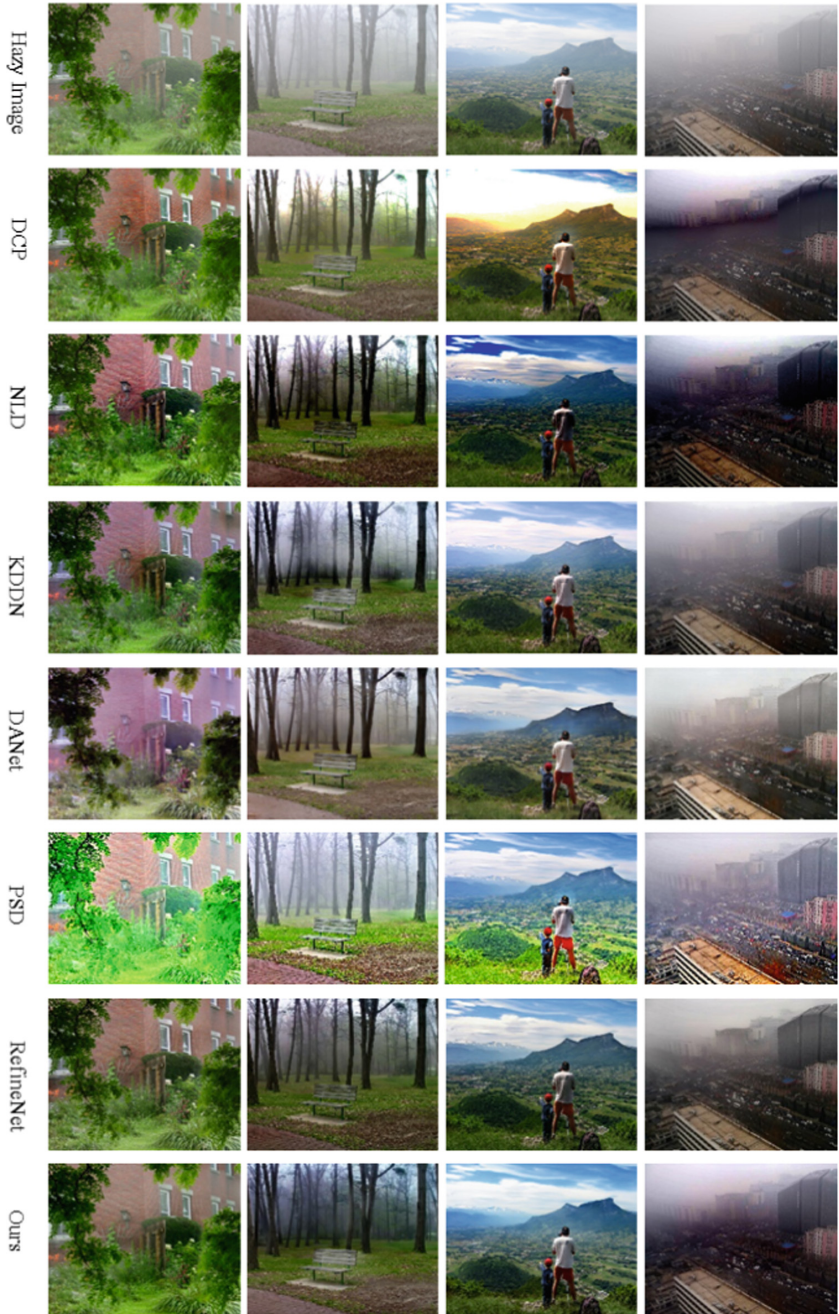
Fig. 4. Comparison of the state-of-the-art dehazing methods on the real-world images.

to guide the student network during pixel-wise comparison. The results are shown in Table 2, it demonstrates that the proposed MGDNet achieves the best performance of image dehazing in terms of PSNR and SSIM. Moreover, adding DCP teacher network, we can improve PSNR from 19.68 dB to 20.32 dB and enhance SSIM by 0.01. And adding NLD teacher network, we can also improve the metrics by 0.56 dB and 0.007. These results show that the combination of prior-based methods improves the performance in outdoor scenes although the network is only trained by indoor images, and the DCP is more efficient than the NLD for our method. Moreover, adding both of DCP teacher network and NLD teacher network also provides with a little gain, which means that combining with multiple priors further improve the generalization ability since unilateral prior cannot hold in various scenes. Additionally, we have also noticed that the addition of PDL drops the performance. It shows that directly using pixel-wise distillation losses may degrade final results since prior-based dehazing images always contain some negative information such as color shifts and artifacts. Fortunately, with the help of DWT, the network can alleviate the distortions in the outputs of teacher networks and further improves the metrics to 20.49 dB and 0.904, respectively.

Table 2. Comparison of variants with different components on the outdoor dataset of SOTS.

Method	PSNR	SSIM
Student	19.68	0.882
Student+DCP	20.32	0.892
Student+NLD	20.24	0.889
Student+DCP+NLD	20.34	0.902
Student+DCP+NLD+PDL	20.24	0.897
Ours	20.49	0.904

5   Conclusion

In this paper, we propose a multi-priors guided dehazing network (MGDNet) based on knowledge distillation, which combines the complementary merits of prior-based dehazing methods and learning-based dehazing methods. Specifically, we pretrain two teacher networks to efficiently use the partially correct features of dark channel prior (DCP) and non-local dehazing (NLD) dehazed results by both of feature-level and pixel-level distillation losses. Moreover, we adopt a high-pass discrete wavelet transform (DWT) before pixel-level guiding to alleviate the negative information of prior dehazed images such as color shifts and artifacts. And the added features of two teacher networks can be refined during the supervised training of student network. Experiments on real-world images demonstrate that the proposed MGDN achieves favorable dehazing effect by both of the quantitative and qualitative comparisons.

References

1. Ancuti, C.O., Ancuti, C.: Single image dehazing by multi-scale fusion. IEEE Trans. Image Process. **2**(8), 3271–3282 (2013)
2. Wang, N., Cui, Z., Su, Y., He, C., Li, A.: Multiscale supervision-guided context aggregation network for single image dehazing. IEEE Signal Process Lett. **29**, 70–74 (2022)
3. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. IEEE Trans. Pattern Anal. Mach. Intell. **33**(12), 2341–2353 (2011)
4. Fattal, R.: Dehazing using color-lines. ACM Trans. Graph. **34**(1), 1–14 (2014)
5. Zhu, Q., Mai, J., Shao, L.: A fast single image haze removal algorithm using color attenuation prior. IEEE Trans. Image Process. **24**(11), 3522–3533 (2015)
6. Berman, D., Treibitz, T., Avidan, S.: Non-local image dehazing. In: Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), LAS Vegas , pp. 1674–1682. IEEE (2016)
7. Cai, B., Xu, X., Jia, K., Qing, C., Tao, D.: DehazeNet: an end-to-end system for single image haze removal. IEEE Trans. Image Process. **25**(11), 5187–5198 (2016)
8. Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., Yang, MH.: Single image dehazing via multi-scale convolutional neural networks. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9906, pp. 154–169. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46475-6_10
9. Li, B., Peng, X., Wang, Z., Xu, J., Feng, D.: AOD-Net: all-in-one dehazing network. In: IEEE International Conference on Computer Vision (ICCV), Venice, pp. 4770–4778. IEEE (2017)
10. Zhang, H., Pattel, V. M.: Densely connected pyramid dehazing network. In: Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, pp. 3194–3203. IEEE (2018)
11. Liu, X., Ma, Y., Shi, Z, Chen, J.: GridDehazeNet: attention-based multi-scale network for image dehazing. In: IEEE International Conference on Computer Vision (ICCV), Seoul, Korea, pp. 7313–7322. IEEE (2019)
12. Qin, X., Wang, Z., Bai, Y., Xie, X. Jia, H.: FFA-Net: feature fusion attention network for single image dehazing. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, pp. 1–8. IEEE (2019)
13. Liu, X., Suganuma, M., Sun, Z., Okatani, T.: Dual residual networks leveraging the potential of paired operations for image restoration. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, pp. 7000–7009. IEEE (2019)
14. Ren, W., et al.: Gated fusion network for single image dehazing. In: Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, pp. 3253–3261. IEEE (2018)
15. Qu Y, Chen Y, Huang J, Xie, Y.: Enhanced Pix2pix dehazing network. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA , pp. 8152–8160. IEEE (2019)
16. Dong H, et al.: Multi-scale boosted dehazing network with dense feature fusion. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, pp. 2154–2164. IEEE (2020)
17. Shao, Y., Li, L., Ren, W., Gao, C., Sang, N.: Domain adaptation for image dehazing. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, pp. 2144–2155. IEEE (2020)
18. Zhao, S., Zhang, L., Shen, Y., Zhou, Y.: RefineDNet: a weakly supervised refinement framework for single image dehazing. IEEE Trans. Image Process. **30**(1), 3391–3404 (2021)
19. Chen, Z., Wang, Y., Yang, Y., Liu, D.: PSD: principled synthetic-to-real dehazing guided by physical priors. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7176–7185 (2021)

20. Hong, M., Xie, Y., Li, C., Qu, Y.: Distilling image dehazing with heterogeneous task imitation. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, pp. 3459–3468. IEEE (2020)
21. Chen, D., et al.: Gated context aggregation network for image dehazing and deraining. In: IEEE Winter Conference on Applications of Computer Vision (WACV), Hilton Waikoloa Village, Hawaii, pp. 1375–1383 (2019)
22. Wang, Z., Ji, S.: Smoothed dilated convolutions for improved dense prediction. In: WIREs Data Mining and Knowledge Discovery, vol. 35, no. 2, pp. 2468–2495 (2021)
23. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, pp. 770–778, IEEE (2016)
24. Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y.: Residual dense network for image super-resolution. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, pp. 2472–2481. IEEE (2018)
25. Huang, G., Liu, Z., Laurens, V., Weinberger, K. Q.: Densely connected convolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), LAS Vegas , pp. 2261–2269 (2016)
26. Zhao, H., Gallo, O., Frosio, I., Kautz, J.: Loss functions for image restoration with neural networks. IEEE Trans. Comput. Imaging **3**(1), 47–57 (2017)
27. Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9906, pp. 694–711. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46475-6_43
28. Fu, M. Liu, H., Yu, Y., Chen, J., Wang, K.: DW-GAN: a discrete wavelet transform GAN for nonhomogenous dehazing. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 203–212 (2021)
29. Li, B., et al.: Benchmarking single-image dehazing and beyond. IEEE Trans. Image Process. **28**(1), 492–505 (2019)
30. Ancuti, C.O., Ancuti, C., Timofte, R., Vleeschouwer, C.D.: I-HAZE: a dehazing benchmark with real hazy and haze-free indoor images. In: IEEE Computer Vision and Pattern Recognition (CVPR), Salt Lake City, pp. 746–754. IEEE (2018)
31. Ancuti, C.O., Ancuti, C., Timofte, R., Vleeschouwer, C.D.: O-HAZE: a dehazing benchmark with real hazy and haze-free outdoor images. In: IEEE Computer Vision and Pattern Recognition (CVPR), Salt Lake City, pp. 754–762. IEEE (2018)
32. Fattal, R.: Single image dehazing. ACM Trans. Graph. **27**(3), 1–9 (2008)
33. Kingma, D., Ba, J.: Adam: a method for stochastic optimization. In: 3rd International Conference on Learning Representations (ICLR), San Diego, CA, USA (2015)