Supplementary Material for PSD: Principled Synthetic-to-Real Dehazing Guided by Physical Priors

Zeyuan Chen[1], Yangchao Wang[2], Yang Yang[2], Dong Liu[1]
[1]University of Science and Technology of China, Hefei, China
[2]University of Electronic Science and Technology of China, Chengdu, China
nnice1216@mail.ustc.edu.cn, yancyycwang@outlook.com, dlyyang@gmail.com, dongeliu@ustc.edu.cn

1. Discussion about Weight Parameters

As mentioned in the paper, the loss function provided by our physical prior committee is

$$\mathcal{L}_{com} = \lambda_d \mathcal{L}_{DCP} + \lambda_b \mathcal{L}_{BCP} + \lambda_c \mathcal{L}_{CLAHE} \tag{1}$$

where $\mathcal{L}_{DCP}$, $\mathcal{L}_{BCP}$, and $\mathcal{L}_{CLAHE}$ are dark channel prior loss, bright channel prior loss, and CLAHE reconstruction loss, respectively. λ_d, λ_b, and λ_c are trade-off weights.

We can see that the choice for loss weights is not a trivial task. Currently, we set them as fixed parameters across the training process. However, it is for sure that different hazy images have a different emphasis and confidence on physical priors, and the best weight settings should be varied based on the input. We had implemented a reinforcement learning technique [2] to adapt the weights based on the input, but the results do not show significant gain than the fixed weights, and the learned weights are similar across different images. We shall remark that the phenomenon may be due the the limited diversity of our training images. In the future, we may explore more optimization techniques for better settings of weight parameters.

2. Details of $\mathcal{L}_{sky}$

We implement $\mathcal{L}_{sky}$ to avoid artifacts and color distortions caused by physical priors. We first estimate the sky region of image using the dark channel prior (DCP). For a restored image $\mathbf{J}$, we calculate its dark channel $\mathcal{J}_{dark}$ by:

$$\mathcal{J}_{dark} = \min_{c \in \{r,g,b\}} \big(\min_{\boldsymbol{y} \in \Omega(\boldsymbol{x})} (J^c(\boldsymbol{y})) \big) \tag{2}$$

where J^c is the c color channel of $\mathbf{J}$ and $\Omega(\boldsymbol{x})$ is a local patch centered at $\boldsymbol{x}$. DCP assumes that the intensity of $\mathcal{J}_{dark}$ is low and tends to be zero, if $\mathbf{J}$ is a haze-free image. However, DCP is typically invalid in sky regions, as pixels in the sky are usually bright and do not have a color channel with very low intensity. Thus, we assume that if the dark channel intensity of a pixel is greater than a threshold λ, it can be approximately treated as a pixel in the sky region. Then we can get the binary mask for the sky region by

$$M_{sky} = \begin{cases} 1, \text{ if } \mathcal{J}_{dark} \geq \lambda \\ 0, \text{ otherwise} \end{cases} \tag{3}$$

Next, we aim to retain the original values of pixels in the sky regions as possible. The loss function $\mathcal{L}_{sky}$ is defined as

$$\mathcal{L}_{sky} = \left\| M_{sky} \odot (J - J_o) \right\|_1 \tag{4}$$

where J and J_o are restored images from the current model $\mathcal{M}$ and the original pre-trained model $\mathcal{M}_o$, respectively.

In practice, λ is set to 150 and the patch size is set to 15×15.

3. Implementation Details of Other Backbones

FFA-Net [11]. In pre-training, the backbone is modified and trained for 20 epochs by the Adam optimizer, with β_1 = 0.9 and β_2 = 0.999. The initial learning rate is set to 10^{-4} and we adopt the same cosine annealing strategy to adjust the learning rate, as in the original paper [11]. In fine-tuning, we train the network for 10 epochs, with an initial learning rate of 10^{-4} that is decayed by 0.5 for every two epochs. The trade-off weights in loss function are set to: $\lambda_d = 10^{-3}$, $\lambda_b = 0.01$, and $\lambda_c = 2$.

GridDehazeNet [10]. In pre-training, the backbone is modified and trained for 10 epochs by the Adam optimizer, with β_1 = 0.9 and β_2 = 0.999. The initial learning rate is set to 10^{-3} with a decay rate of 0.5 for every 2 epochs. In fine-tuning, we train the network for 10 epochs, with an initial learning rate set to 10^{-4} that is decayed by 0.5 for every two epochs. The trade-off weights in the loss function are set to: $\lambda_d = 5 \times 10^{-3}$, $\lambda_b = 0.01$, and $\lambda_c = 1$.

4. Additional Ablation Studies

We conduct two more ablation studies on the learning without forgetting (LwF) loss and the sky region loss, respectively.

(a) w/o LwF Loss (b) w/ LwF Loss

Figure 1: Comparison for LwF Loss.

(a) w/o sky region loss (b) w/ sky region loss

Figure 2: Comparison for sky region loss.

4.1. LwF Loss

LwF loss can be viewed as a regularization term during the fine-tuning, which prevents the model from over-fitting to the small-scale training data (of real hazy images). The visual comparison is presented in Fig. 1. Without LwF loss, the dehazing result suffers from color distortion and over-saturation, while the result with LwF loss is of good color distribution.

4.2. Sky Region Loss

Sky region loss is proposed for the circumstances when the physical priors fail to handle the sky regions well. As shown in Fig. 2, the result without sky region loss suffers from noticeable artifacts in the sky, while the result with sky region loss is more realistic and visually pleasing.

5. Additional Experimental Results

More Comparison Results. In Figs. 3–11, we compare more dehazing results of PSD with state-of-the-art dehazing methods. All the results of PSD are conducted on the MS-BDN [3] backbone. Hazy images in Figs. 3–8 are picked from RTTS and HSTS (both are subsets of RESIDE [8]), and images in Figs. 9–11 are from authors of previous work [4–6]. We can observe that PSD restores more image details and generates cleaner results than other methods.

More Results for Other Backbones. We show more experimental results of PSD upon different backbones to illustrate that PSD is generally applicable. All the images are picked from URHI and they are excluded from the training data. As shown in Figs. 12–18, PSD produces significant improvements on backbone models and performs well against the state-of-the-art domain adaptation dehazing [13]. For simplicity of notation, we use PSD (MSBDN) to represent the model modified and trained by PSD on the MSBDN backbone. Similarly, we have PSD (FFA) and PSD (Grid).

References

[1] Dana Berman, Tali Treibitz, and Shai Avidan. Non-local image dehazing. In *CVPR*, pages 1674–1682, 2016. 3, 4, 5, 6, 7, 8, 9, 10, 11

[2] Wuyang Chen, Zhiding Yu, Zhangyang Wang, and Anima Anandkumar. Automated synthetic-to-real generalization. In *Proceedings of Machine Learning and Systems 2020*, pages 8272–8282, 2020. 1

[3] Hang Dong, Jinshan Pan, Lei Xiang, Zhe Hu, Xinyi Zhang, Fei Wang, and Ming-Hsuan Yang. Multi-scale boosted dehazing network with dense feature fusion. In *CVPR*, pages 2157–2167, 2020. 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14

[4] Raanan Fattal. Single image dehazing. *ACM Transactions on Graphics*, 27(3):1–9, 2008. 2, 9, 10, 11

[5] Raanan Fattal. Dehazing using color-lines. *ACM Transactions on Graphics*, 34(1):1–14, 2014. 2, 9, 10, 11

[6] Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12):2341–2353, 2010. 2, 9, 10, 11

[7] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. AOD-Net: All-in-one dehazing network. In *ICCV*, pages 4770–4778, 2017. 3, 4, 5, 6, 7, 8, 9, 10, 11

[8] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing*, 28(1):492–505, 2018. 2

[9] Lerenhan Li, Yunlong Dong, Wenqi Ren, Jinshan Pan, Changxin Gao, Nong Sang, and Ming-Hsuan Yang. Semi-supervised image dehazing. *IEEE Transactions on Image Processing*, 29:2766–2779, 2020. 3, 4, 5, 6, 7, 8, 9, 10, 11

[10] Xiaohong Liu, Yongrui Ma, Zhihao Shi, and Jun Chen. Grid-DehazeNet: Attention-based multi-scale network for image dehazing. In *ICCV*, pages 7314–7323, 2019. 1, 12, 13, 14

[11] Xu Qin, Zhilin Wang, Yuanchao Bai, Xiaodong Xie, and Huizhu Jia. FFA-Net: Feature fusion attention network for single image dehazing. In *AAAI*, pages 11908–11915, 2020. 1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14

[12] Yanyun Qu, Yizi Chen, Jingying Huang, and Yuan Xie. Enhanced pix2pix dehazing network. In *CVPR*, pages 8160–8168, 2019. 3, 4, 5, 6, 7, 8, 9, 10, 11

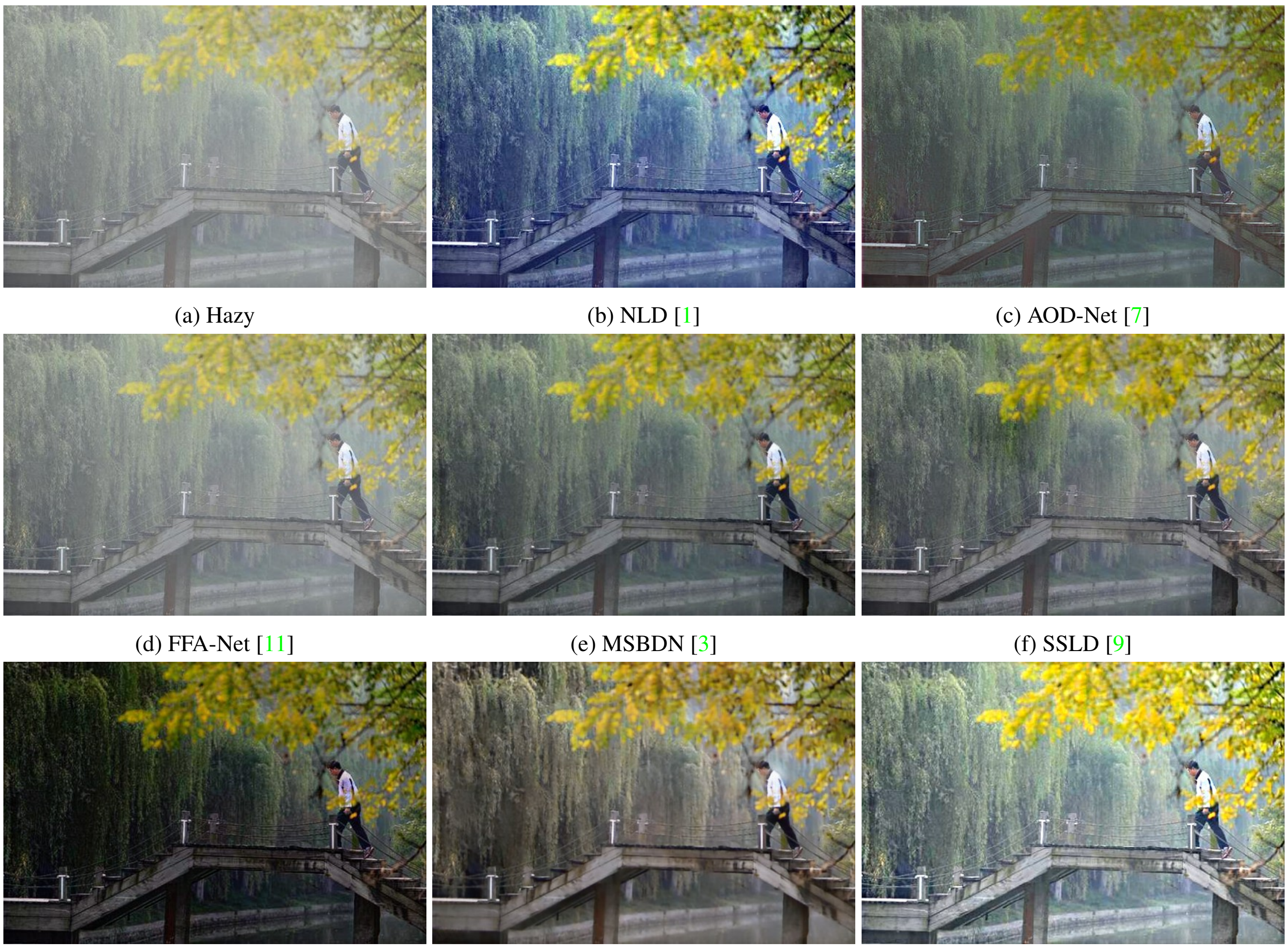
[13] Yuanjie Shao, Lerenhan Li, Wenqi Ren, Changxin Gao, and Nong Sang. Domain adaptation for image dehazing. In *CVPR*, pages 2808–2817, 2020. 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14

(a) Hazy (b) NLD [1] (c) AOD-Net [7]

(d) FFA-Net [11] (e) MSBDN [3] (f) SSLD [9]

(g) EPDN [12] (h) DAD [13] (i) PSD

Figure 3: Visual comparison on a real hazy image from RTTS.



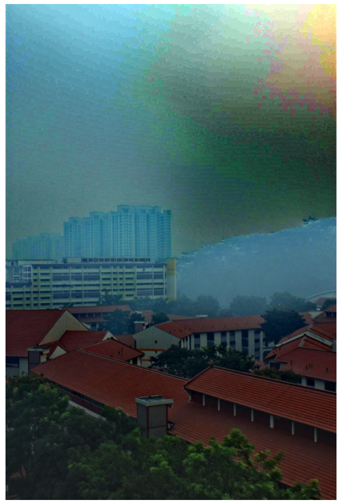
(a) Hazy (b) NLD [1] (c) AOD-Net [7]

(d) FFA-Net [11] (e) MSBDN [3] (f) SSLD [9]

(g) EPDN [12] (h) DAD [13] (i) PSD

Figure 4: Visual comparison on a real hazy image from RTTS.

(a) Hazy



(b) NLD [1]

(c) AOD-Net [7]

(d) FFA-Net [11]

(e) MSBDN [3]

(f) SSLD [9]

(g) EPDN [12]

(h) DAD [13]

(i) PSD

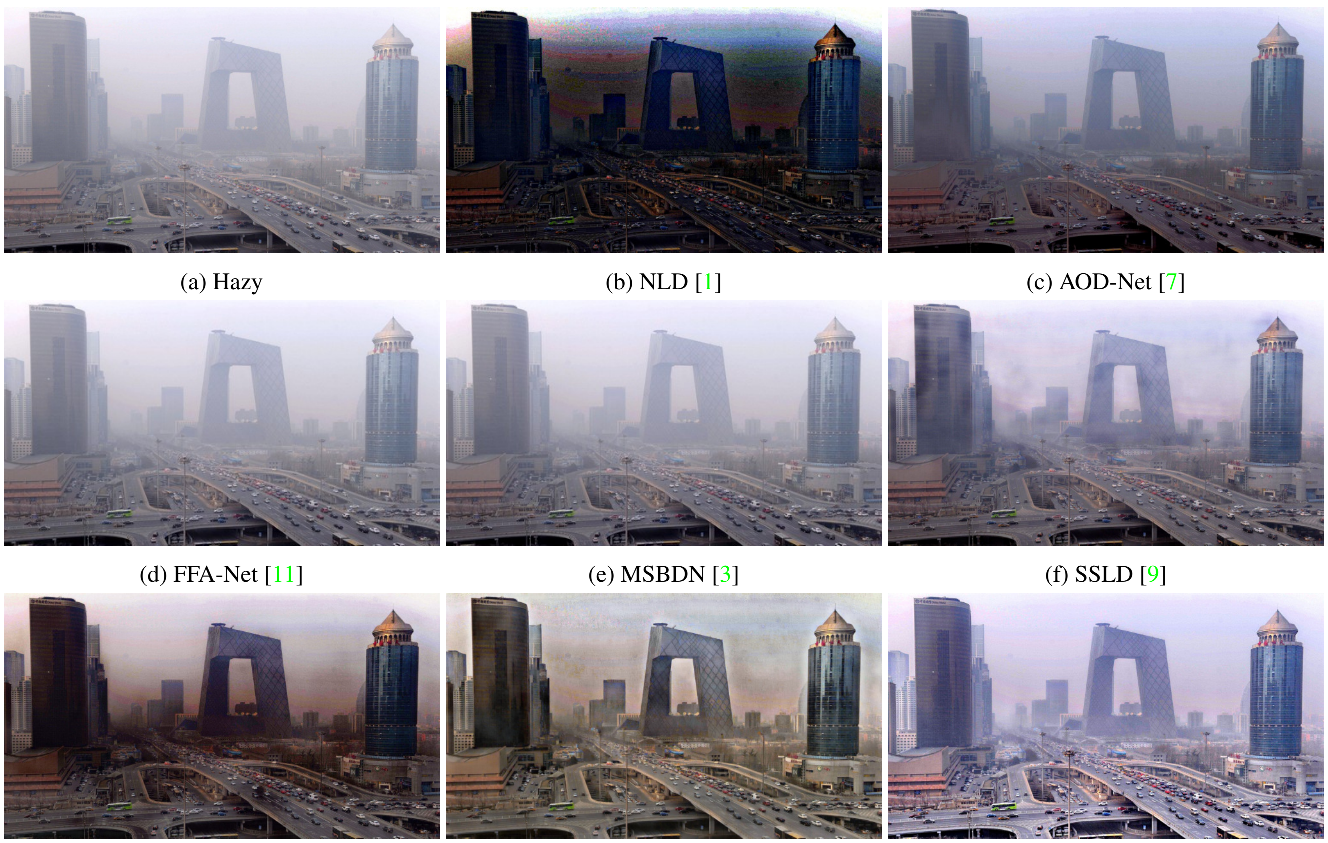
Figure 5: Visual comparison on a real hazy image from HSTS.

(a) Hazy (b) NLD [1] (c) AOD-Net [7]

(d) FFA-Net [11] (e) MSBDN [3] (f) SSLD [9]

(g) EPDN [12] (h) DAD [13] (i) PSD

Figure 6: Visual comparison on a real hazy image from HSTS.



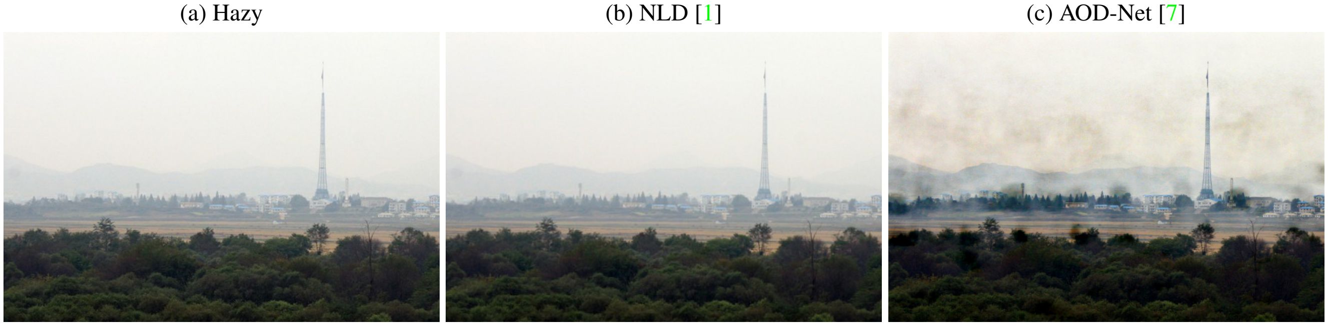
(a) Hazy (b) NLD [1] (c) AOD-Net [7]

(d) FFA-Net [11] (e) MSBDN [3] (f) SSLD [9]

(g) EPDN [12] (h) DAD [13] (i) PSD

Figure 7: Visual comparison on a real hazy image from HSTS.

(a) Hazy (b) NLD [1] (c) AOD-Net [7]



(d) FFA-Net [11] (e) MSBDN [3] (f) SSLD [9]

(g) EPDN [12] (h) DAD [13] (i) PSD

Figure 8: Visual comparison on a real hazy image from HSTS.

(a) Hazy

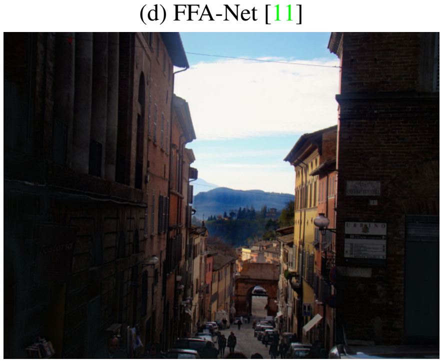
(b) NLD [1]

(c) AOD-Net [7]

(d) FFA-Net [11]

(e) MSBDN [3]

(f) SSLD [9]



(g) EPDN [12]

(h) DAD [13]

(i) PSD

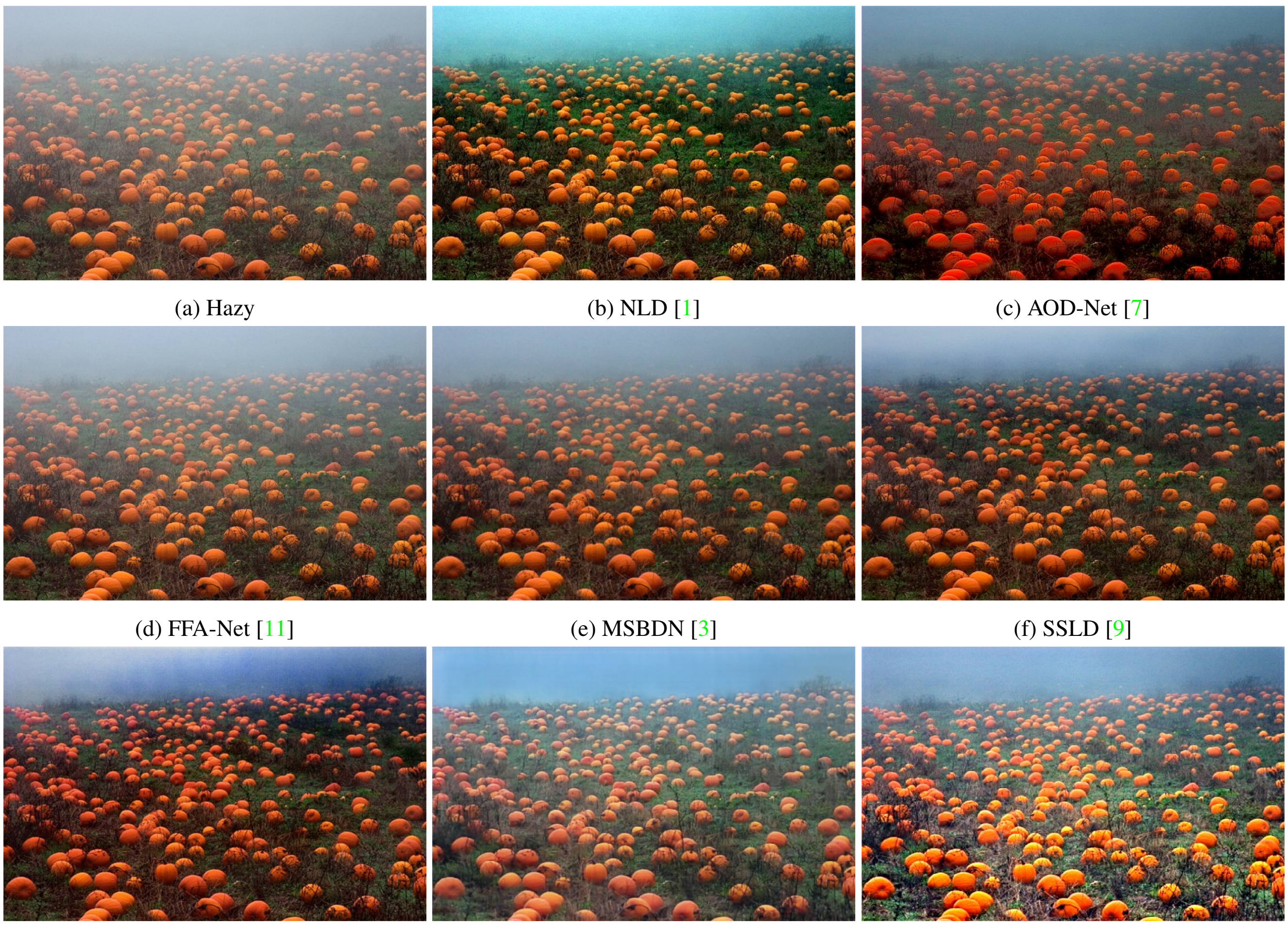
Figure 9: Visual comparison on a real hazy image released from authors of previous work [4–6].

(a) Hazy (b) NLD [1] (c) AOD-Net [7]

(d) FFA-Net [11] (e) MSBDN [3] (f) SSLD [9]

(g) EPDN [12] (h) DAD [13] (i) PSD

Figure 10: Visual comparison on a real hazy image released from authors of previous work [4–6].



(a) Hazy (b) NLD [1] (c) AOD-Net [7]

(d) FFA-Net [11] (e) MSBDN [3] (f) SSLD [9]

(g) EPDN [12] (h) DAD [13] (i) PSD

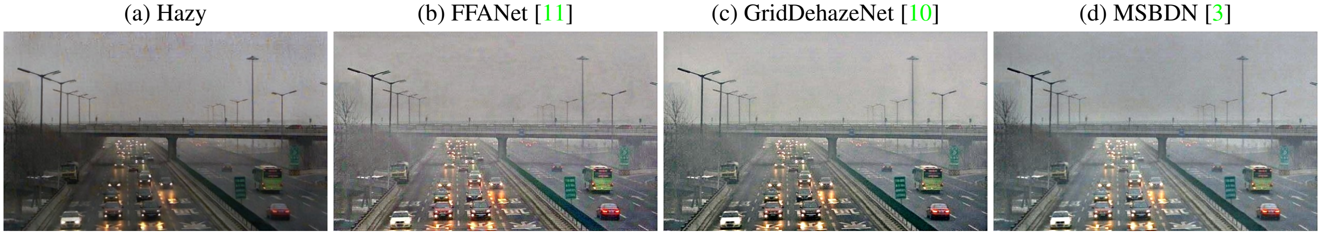
Figure 11: Visual comparison on a real hazy image released from authors of previous work [4–6].

(a) Hazy (b) FFANet [11] (c) GridDehazeNet [10] (d) MSBDN [3]

(e) DAD [13] (f) PSD (FFA) (g) PSD (Grid) (h) PSD (MSBDN)

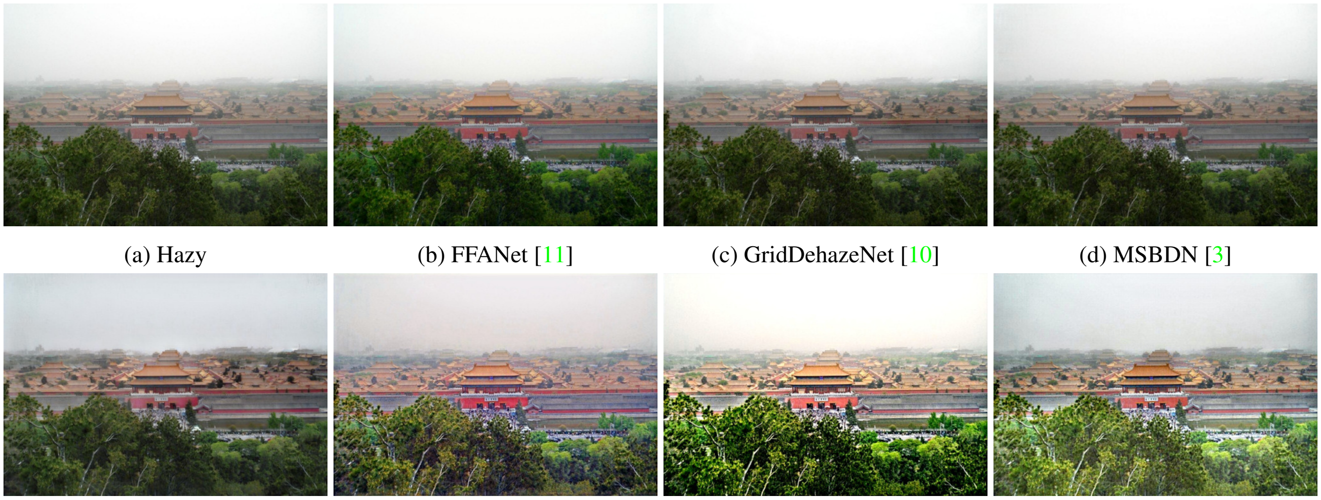
Figure 12: Results of using different backbones.

(a) Hazy (b) FFANet [11] (c) GridDehazeNet [10] (d) MSBDN [3]



(e) DAD [13] (f) PSD (FFA) (g) PSD (Grid) (h) PSD (MSBDN)

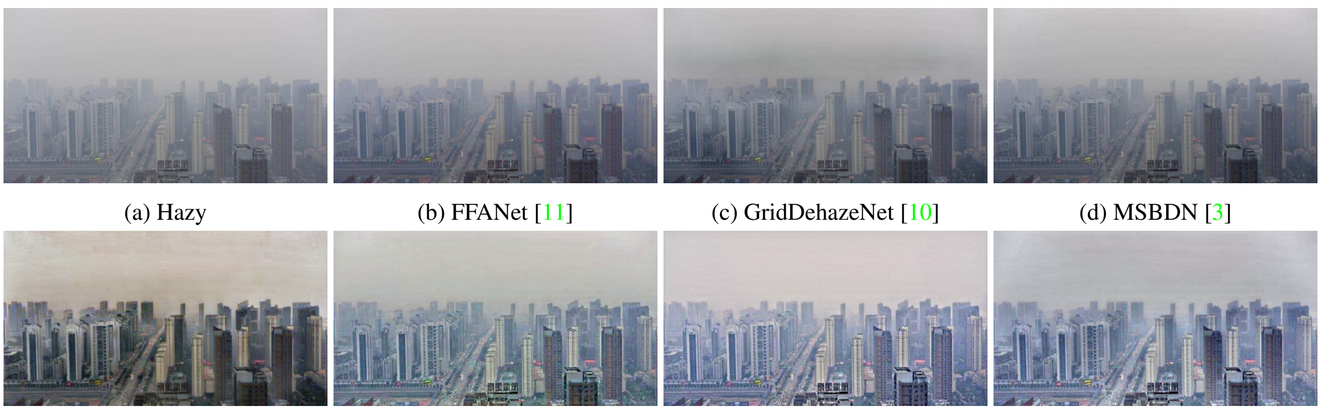
Figure 13: Results of using different backbones.



(a) Hazy (b) FFANet [11] (c) GridDehazeNet [10] (d) MSBDN [3]

(e) DAD [13] (f) PSD (FFA) (g) PSD (Grid) (h) PSD (MSBDN)

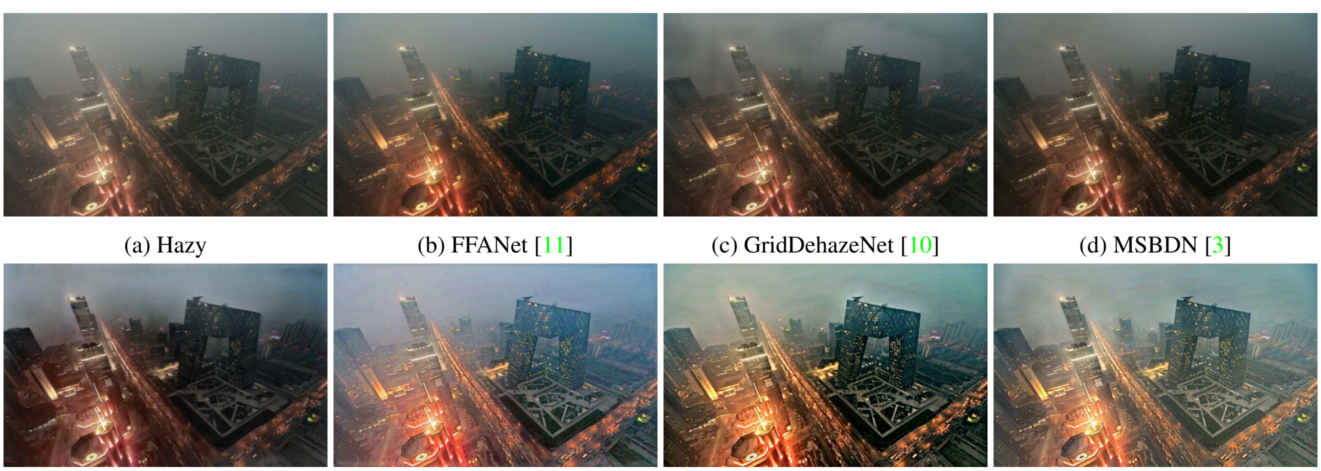
Figure 14: Results of using different backbones.



(a) Hazy (b) FFANet [11] (c) GridDehazeNet [10] (d) MSBDN [3]

(e) DAD [13] (f) PSD (FFA) (g) PSD (Grid) (h) PSD (MSBDN)

Figure 15: Results of using different backbones.



(a) Hazy (b) FFANet [11] (c) GridDehazeNet [10] (d) MSBDN [3]

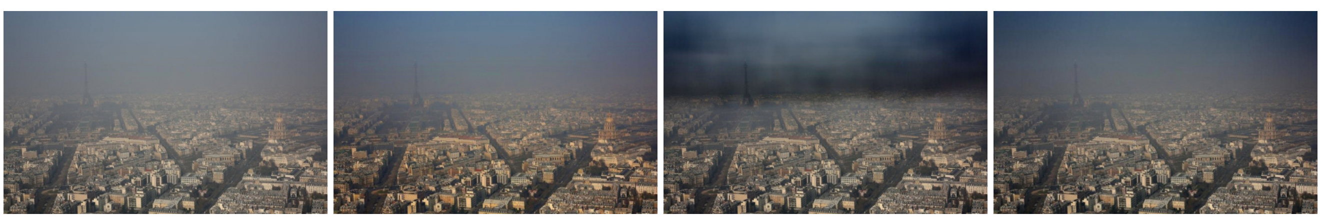
(e) DAD [13] (f) PSD (FFA) (g) PSD (Grid) (h) PSD (MSBDN)

Figure 16: Results of using different backbones.

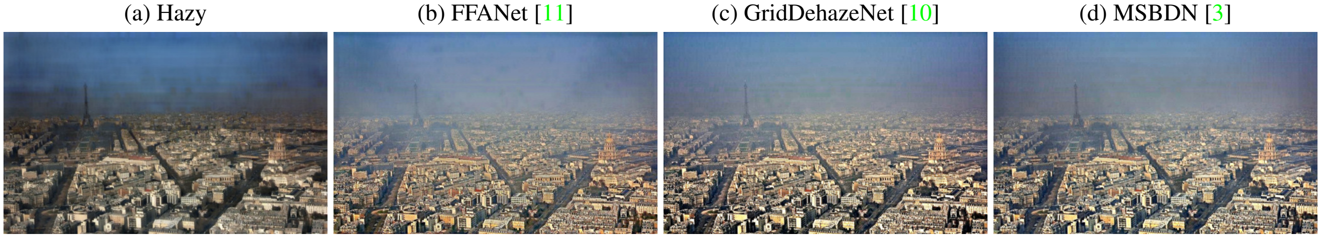
(a) Hazy (b) FFANet [11] (c) GridDehazeNet [10] (d) MSBDN [3]

(e) DAD [13] (f) PSD (FFA) (g) PSD (Grid) (h) PSD (MSBDN)

Figure 17: Results of using different backbones.



(a) Hazy (b) FFANet [11] (c) GridDehazeNet [10] (d) MSBDN [3]



(e) DAD [13] (f) PSD (FFA) (g) PSD (Grid) (h) PSD (MSBDN)

Figure 18: Results of using different backbones.