Physically Disentangled Intra- and Inter-domain Adaptation for Varicolored Haze Removal

Yi Li, Yi Chang*, Yan Gao, Changfeng Yu, Luxin Yan
National Key Laboratory of Science and Technology on Multispectral Information Processing,
School of Artificial Intelligence and Automation, Huazhong University of Science and Technology
{li_yi, yichang, gaoyan_117, ycf, yanluxin}@hust.edu.cn

Abstract

Learning-based image dehazing methods have achieved marvelous progress during the past few years. On one hand, most approaches heavily rely on synthetic data and may face difficulties to generalize well in real scenes, due to the huge domain gap between synthetic and real images. On the other hand, very few works have considered the varicolored haze, caused by chromatic casts in real scenes. In this work, our goal is to handle the new task: real-world varicolored haze removal. To this end, we propose a physically disentangled joint intra- and inter-domain adaptation paradigm, in which intra-domain adaptation focuses on color correction and inter-domain procedure transfers knowledge between synthetic and real domains. We first learn to physically disentangle haze images into three components complying with the scattering model: background, transmission map, and atmospheric light. Since haze color is determined by atmospheric light, we perform intra-domain adaptation by specifically translating atmospheric light from varicolored space to unified color-balanced space, and then reconstructing color-balanced haze image through the scattering model. Consequently, we perform inter-domain adaptation between the synthetic and real images by mutually exchanging the background and other two components. Then we can reconstruct both identity and domain-translated haze images with self-consistency and adversarial loss. Extensive experiments demonstrate the superiority of the proposed method over the state-of-the-art for real varicolored image dehazing.

1. Introduction

Haze, as a common weather phenomenon, would result in low contrast and severe visibility degradation, which not only leads to poor visual quality but also does serious harm to high-level vision tasks, such as scene classification [35], object detection [25] and semantic segmentation [41]. The

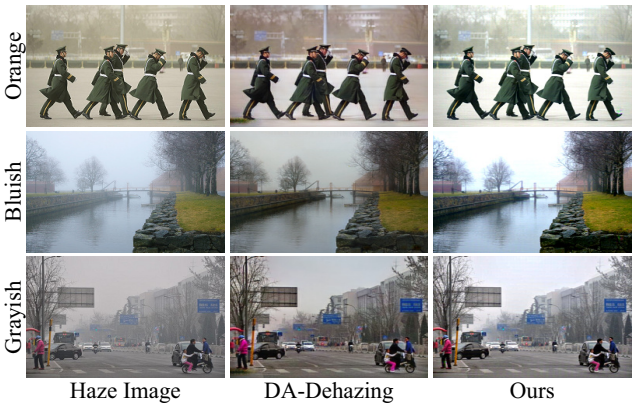

Figure 1. The visual examples of dehazing results for real-world varicolored haze images. The second and third column show the results of DA-Dehazing [42] and proposed method, respectively.

haze procedure can be mathematically formulated via the well-known atmosphere scattering model [29, 34]:

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

where $I(x)$ is observed haze image, and $J(x)$ is haze-free background to be restored. A and $t(x) = e^{-\beta d(x)}$ denote atmospheric light and transmission map, and β and $d(x)$ represent scattering coefficient and depth respectively. The goal of dehazing is to estimate $J(x)$ from hazy input $I(x)$.

The image dehazing has been extensively studied during the past few years [6, 7, 11, 18, 25, 28, 31, 39, 50]. Benefitting from powerful representation ability of convolutional neural networks (CNNs), CNN-based supervised learning methods have made marvelous progress [8, 12, 13, 32, 37, 45]. Since it is difficult to obtain labeled clean groundtruth, it is widely accepted to simulate paired clean/haze images via the scattering model for CNN training. However, the inter-domain gap between real and synthetic data makes these supervised methods hard to generalize well for the real haze images. To overcome this issue, the semi-supervised [9, 27, 33, 42] and the unsupervised dehazing methods [17, 23, 24, 30, 46, 47] have been naturally proposed. The unsupervised methods mainly utilize the prior knowledge of the real images by min-

*Corresponding author.

imizing the unsupervised loss function, such as dark channel prior [18]. The semi-supervised methods not only utilize real images but also incorporate synthetic images for network training. The researchers [9, 27, 33, 42] mainly adopt the domain adaptation technique to transfer the knowledge from synthetic domain to the real haze domain. Although these methods have greatly alleviated the inter-domain shift issue, few of them have considered the intra-domain issue, namely the varicolored haze space in real-world scenes.

The varicolored haze is mainly caused by various particles with different sizes and scattering characteristics [14], as shown in Fig. 1. Most of the previous methods employ the hand-crafted white balance and contrast enhancement techniques to eliminate the color distortion [1, 2, 19]. Recently, the learning-based varicolored image dehazing methods have been proposed [14, 40]. For example, Dudhane *et al.* [14] proposed a color correction module based on gray world assumption [38] to better estimate accurate illumination and restore color distortion. However, these supervised methods mainly utilize synthetic datasets for training, which may face difficulties in handling varicolored haze in real scenes. To our knowledge, we are the first to simultaneously consider the intra-domain gap within varicolored haze spaces and the inter-domain gap between synthetic and real images.

In this work, we formulate the challenging varicolored real hazy removal task into two easier yet physical meaning sub-problems and propose the physically disentangled joint intra- and inter-domain adaptation (**PDI2A**) framework for semi-supervised varicolored dehazing, as shown in Fig. 2. Our model consists of three parts: two intra-domain adaptations for color correction on both real and synthetic inputs, and an inter-domain adaptation for cycle synthetic↔real translation. Specifically, in intra-domain adaptation, we disentangle haze image via scattering model and align color in simpler scalar space, not directly in complex image space. In inter-domain adaptation, disentanglement strategy is also employed to ease translation difficulty. Both physical- and learning-based reconstruction are designed to exchange knowledge between synthetic&real data for semi-supervised learning. Moreover, we apply self-consistency and adversarial loss for better information preserving and translation. We summarize the main contributions as follows:

- Our work focuses on a novel yet practical task: varicolored image dehazing in real scenes. Compared with existing methods which consider either intra-domain (varicolored spaces) or inter-domain gap (synthetic and real spaces), we are the first to joint consider both gaps simultaneously.
- We propose a semi-supervised joint intra- and inter-domain adaptation framework, which divides the challenging varicolored real image dehazing into two easier yet physical meaning subproblems: intra-domain color alignment and inter-domain synthetic-to-real haze removal, significantly easing the learning procedure.
- We propose to resolve the color/hazy artifacts in the simpler disentangled space, not the complex image space, in which the consistent background is well preserved and only the critical haze characteristics are transferred, significantly easing the translation difficulty. Extensive experiments on real and synthetic data demonstrate the superiority of our method both quantitatively and qualitatively.

2. Related Work

2.1. Varicolored Haze Removal

Varicolored haze would lead to not only limited visibility but also severe color distortion in real scenes. To eliminate haze effects, numerous methods have been proposed to remove haze via efficient priors [11, 15, 16, 18, 43] or powerful CNNs [7, 8, 13, 25, 32, 39, 40, 45]. However, most existing methods mainly focus on improving visibility and neglect the color distortion caused by varicolored haze.

Recently, color correction has been taken into consideration for practical applications such as sandstorms and smog. Most of these methods utilized white balance technique to rectify the color distribution first, followed by visibility enhancement to generate haze-free images [1, 2, 19, 40]. Ren *et al.* [40] derived color balanced input through gray world assumption (GWA) [38] along with contrast-enhanced one to generate haze-free image through gated fusion technique. Dudhane *et al.* [14] proposed a cascaded network, including the GWA-based haze color correction module and visibility improvement module. Nevertheless, these methods heavily rely on synthetic dataset and may face difficulties to generalize well in real scenes due to the huge gap between synthetic and real data [42]. In this work, we solve varicolored real haze removal through a semi-supervised paradigm.

2.2. Domain Adaptation

Domain adaptation aims at aligning the distribution shift between source and target domains. As for image dehazing tasks, several works have made use of domain adaptation from feature- or pixel-level to solve the domain shift between synthetic and real data [9, 27, 33, 42]. For example, Li *et al.* [27] introduced a semi-supervised dehazing method, in which a weight-shared network learned domain invariant features for both synthetic and real data. Shao *et al.* [42] performed pixel-wise image translation through CycleGAN [48] to bridge the gap between synthetic and real domains, followed by two restoration networks specially for synthetic and real image dehazing. Although considerable progress has been made for real haze removal, these methods mainly consider aligning inter-domain gap between synthetic&real data while neglecting intra-domain gap caused by color distortions. In this work, we jointly consider both intra- and inter-domain gaps, in which intra-domain adaptation is performed to rectify color distortion, hence easing difficulties

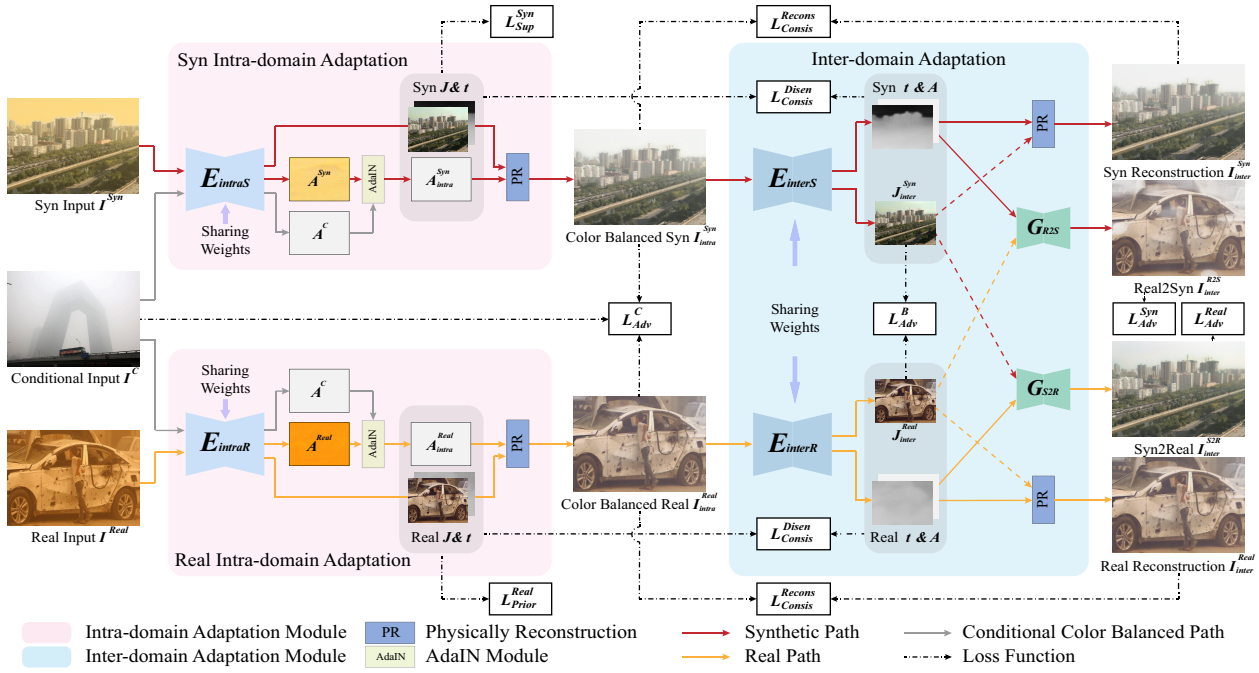

Figure 2. Architecture of physically disentangled intra- and inter-domain adaptation framework for varicolored haze removal in real scene. The whole framework is composed of two intra-domain adaptation modules for color correction and one inter-domain adaptation module for synthetic↔real translations. Intra-domain adaptation firstly performs color correction to ease difficulties of translations between synthetic&real domains. Moreover, instead of performing image-to-image translation directly, we physically disentangle the image according to scattering model and only translate critical haze components while keeping identity background preserved, leading to accurate translation.

of latter inter-domain adaptation among synthetic&real data.

2.3. Image Translation

Recently, amounts of works have been proposed for image translation. These methods usually translate images from one domain to another with pixel- or feature-level alignment. For example, Zhu *et al.* [48] proposed CycleGAN to perform bidirectional translation between two domains via both adversarial and cycle consistency losses. Recently, to further improve variance, disentangled representations have been proposed for multi-modal image translation. MUINT [21] and DRIT [22] decoupled image content and style so as to multi-modal translation, in which content and style served as latent spaces aligned only via adversarial loss. In this work, we propose to endow disentanglement with explicit physical meaning through scattering model and perform translation exactly on critical haze components while keeping identity background unchanged, thus easing translation difficulties.

3. Proposed Method

3.1. Framework Overview

Given a haze image I, varicolored image dehazing aims at estimating latent image J. The key problem is the huge gaps exist not only between synthetic training and real testing data (inter-domain) but also among varicolored haze (intra-domain). Existing methods either try to bridge intra- or inter-domain gap, seldom works consider them simultaneously.

To bridge both gaps simultaneously, we propose to jointly consider both intra- and inter-domain adaptation within a unified physically disentangled image translation framework. The core idea of the proposed method is to first get rid of color distortion and then transfer knowledge between synthetic&real domain. Specifically, as shown in Fig. 2, the proposed PDI2A consists of two intra-domain adaptation modules and one inter-domain adaptation module. Two intra-domain adaptation take varicolored haze images I^{Syn}, I^{Real} as input and perform varicolor→color-balanced translation to acquire I^{Syn}_{intra} and I^{Real}_{intra}. And then inter-domain adaptation performs not only cross-domain translation for I^{S2R}_{inter} and I^{R2S}_{inter} but also self-reconstruction for I^{Syn}_{inter} and I^{Real}_{inter}. Moreover, we construct consistency between intra-domain and inter-domain via self-consistency. Overall, such a progressive multi-stage strategy would significantly ease difficulties of multi-domain translations.

3.2. Physically Disentangled Translation

Intra-domain: atmospheric light oriented color correction. According to Eq. (1), haze images I can be decomposed into three components: clear background J, transmission t, and atmospheric light A. The first term represents scene radiance, while the second and third terms denote the

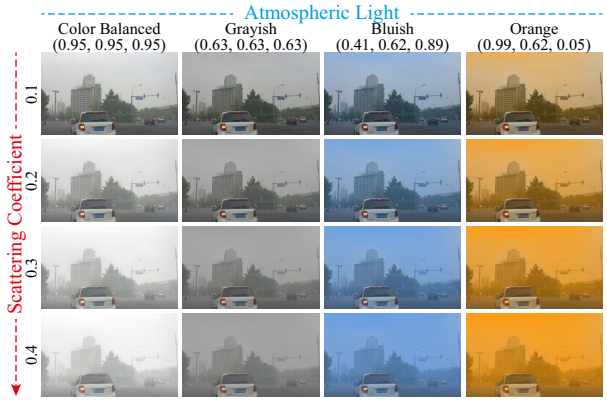

Figure 3. Analysis of the influence of different parameters in atmosphere scattering model. Each row and column represents different parameters of atmospheric light and scattering coefficient. The scattering coefficient mainly controls the density of haze, while atmospheric light is responsible for different color distortions.

effect of haze and corresponding color cast [14]. *That is to say, atmospheric light A determines color distortion degree*.

To illustrate this, we visualize how the atmospheric light A influences the color distortion. In Fig. 3, we synthesize haze images with diverse atmosphere lights A and scattering coefficients β respectively. Along the horizontal axis, we gradually change the atmospheric light A with different values with other two components fixed. Along the vertical axis, we gradually increase the scattering coefficient β with other two components fixed. We can observe that the scattering coefficient β mainly controls the density of haze while atmosphere light A dominates the color cast. This observation motivates us to perform color correction, namely the intra-domain translation only on atmospheric light space, instead of directly on the haze images.

Consequently, our intra-domain adaptation module first physically disentangle varicolored haze image into J, A, and t via multi-head encoder ($E_{intra} : I \rightarrow \{J, A, t\}$). Then, we can easily align the color in atmosphere light space via classical AdaIN module [20] with estimated atmospheric light A and the color-balanced one A^C serving as source and target style respectively, translating the varicolored A to the color-balanced A_{intra}. Finally, A_{intra} is further combined to reconstruct color-balanced haze image I_{intra} through Eq. (1). Note that standard A^C is hard to preset in advance and fixed one may lead to model collapse. Thus, we randomly select one real color-free hazy image I^C to generate diverse and realistic atmospheric light each time. Compared with previous image translation methods, the proposed PDI2A performs translation only on disentangled color-relevant component while keeping background and transmission consistent, greatly ensuring the color correction performance.

Inter-domain: background preserved cross-domain translation. The inter-domain adaptation is to transfer the labeled knowledge from the synthetic haze image to the unlabelled real haze image. We further learn to disentangle the color-balanced syn/haze images into three components $(J_{inter}, A_{inter}, t_{inter})$. The key insight of our inter-domain adaptation is that clean background from synthetic and real domains lie on the same distributions, while corresponding haze images lie on different distributions. This motivates us to take image background as domain-invariant content space, while transmission and atmospheric light as domain-specific attribute space. Then we could perform the cross-domain translation by exchanging the attribute spaces of synthetic and real images. Meanwhile, we could also reconstruct the identical haze image with the disentangled components. Since transmission t is highly correlated with background J, physical reconstruction (PR) via atmospheric model is employed when transmission t and image J are matched and data-driven learning generator is designed when mismatched. Moreover, the reconstruction and adversarial losses are jointly performed to ensure satisfying knowledge preserving and translation. Finally, we build the disentangle consistency between the intra- and inter-domain translation to further facilitate the learning of disentanglement. In contrast to the existing image translation techniques [22, 48], the proposed method embeds explicit physical meaning to disentangle process, and keep the background J consistent while translating only haze relevant components, leading to better disentanglement and translation performance.

3.3. Loss Function

In this section, we give descriptions about loss functions in PDI2A, which can be roughly divided into three classes: adversarial loss, consistency loss, and disentangle loss.

Adversarial Loss. Firstly, to enforce color correction, we employ a color discriminator D_{Color} which attempts to distinguish translated haze image I^{Syn}_{intra} and I^{Real}_{intra} from original color balanced one I^C. And network tries to fool discriminator by generating more realistic color-balanced haze images. Thus color adversarial loss can be formulated as:

$$\begin{aligned}\mathcal{L}^C_{Adv} = &\mathbb{E}_{I^{Syn}_{intra}}[log(1 - D_C(I^{Syn}_{intra}))]+ \\ &\mathbb{E}_{I^{Real}_{intra}}[log(1 - D_C(I^{Real}_{intra}))] + \alpha\mathbb{E}_{I^C}[logD_C(I^C)],\end{aligned} \tag{2}$$

where α is the balance weight. Then, since we assume disentangled clean background share the same distribution, we adopt background discriminator D_B to distinguish weather haze exists in J^{Syn}_{inter} and J^{Real}_{inter}:

$$\begin{aligned}\mathcal{L}^B_{Adv} = &\mathbb{E}_{J^{Syn}_{inter}}[log(1 - D_B(J^{Syn}_{inter}))]+ \\ &\mathbb{E}_{J^{Real}_{inter}}[log(1 - D_B(J^{Real}_{inter}))] + \alpha\mathbb{E}_{J^C}[logD_B(J^C)],\end{aligned} \tag{3}$$

where J^C is images from haze-free domain. Furthermore, for generated synthetic and real haze images, we apply two discriminators D_S, D_R to produce adversarial loss. For example, the translated synthetic image I^{S2R}_{inter} should share the same distribution with the synthetic color balanced haze images $\widetilde{I}^{Real}$. Thus the loss function can be formulated as:

$$\mathcal{L}_{Adv}^{Real} = \mathbb{E}_{I_{inter}^{S2R}}[log(1 - D_R(I_{inter}^{S2R}))] + \mathbb{E}_{\widetilde{I}^{Real}}[logD_R(\widetilde{I}^{Real})]. \quad (4)$$

Similar adversarial loss $\mathcal{L}_{Adv}^{Syn}$ is also imposed to ensure the synthetic haze image I_{inter}^{R2S} generation.

Consistency Loss. Since the disentangle procedure may unexpectedly degrade the information, we employ the consistency loss both in the haze image domain and disentangled components domain, serving as self-supervision of the disentangle procedure. Firstly, in inter-domain adaptation module, haze images are not only performed inter-domain translation but also performed self-reconstruction from disentangled components. We define reconstruction consistency loss as:

$$\mathcal{L}_{Consis}^{Recons} = \|I_{intra}^{Real} - I_{inter}^{Real}\|_1 + \|I_{intra}^{Syn} - I_{inter}^{Syn}\|_1. \quad (5)$$

Moreover, since disentangle and entangle are mutually inverse processes, the disentangled components in intra- and inter-domain adaptation module should be consistent. Thus we define the disentangle consistency loss as:

$$\begin{aligned}\mathcal{L}_{Consis}^{Disen} = &\|J_{intra}^{Real} - J_{inter}^{Real}\|_1 + \|J_{intra}^{Syn} - J_{inter}^{Syn}\|_1 + \\ &\|t_{intra}^{Real} - t_{inter}^{Real}\|_1 + \|t_{intra}^{Syn} - t_{inter}^{Syn}\|_1 + \\ &\|A_{intra}^{Real} - A_{inter}^{Real}\|_1 + \|A_{intra}^{Syn} - A_{inter}^{Syn}\|_1.\end{aligned} \quad (6)$$

Disentangle Loss. In addition to adversarial losses and consistency losses, for synthetic haze images which have groundtruth for each disentangled component, we utilize L1 loss to supervise training of synthetic path:

$$\begin{aligned}\mathcal{L}_{Sup}^{Syn} = &\|J_{intra}^{Syn} - J_{GT}^{Syn}\|_1 + \\ &\|t_{intra}^{Syn} - J_{GT}^{Syn}\|_1 + \|A^{Syn} - A_{GT}^{Syn}\|_1.\end{aligned} \quad (7)$$

Moreover, we also utilize efficient prior loss for real path training such as dark channel to enforce disentanglement:

$$\mathcal{L}_{Prior}^{Real} = \| \min_{y \in N(x)} [\min_{c \in r,g,b} (J_{intra}^{Real}(y,c))] \|_1, \quad (8)$$

where x and y are pixel coordinates and c represents the color channel of estimated background image J_{intra}^{Real}. $N(x)$ denotes the local neighborhood centered at x.

Overall Objective Function. The full objectivfe function contains three losses as follow:

$$\begin{aligned}\mathcal{L}_{Total} = &\lambda_{Adv}(\mathcal{L}_{Adv}^{C} + \mathcal{L}_{Adv}^{B} + \mathcal{L}_{Adv}^{Real} + \mathcal{L}_{Adv}^{Syn}) + \\ &\lambda_{Consis}(\mathcal{L}_{Consis}^{Recons} + \mathcal{L}_{Consis}^{Disen}) + \lambda_{Disen}(\mathcal{L}_{Sup}^{Syn} + \mathcal{L}_{Prior}^{Real}).\end{aligned} \quad (9)$$

Through the full objective function, we alleviate the problem of lacking paired haze and haze-free images. In our framework, both intra- and inter-domain adaptation are constrained by several supervised losses and consistency losses to guarantee disentanglement and translation performance.

3.4. Implementation Details

The whole PDI2A framework is implemented on Pytorch with a single RTX 3090 GPU. We utilize modified U-Net as disentangle network E_{intra} and E_{inter} in both intra- and inter-domain adaptations. We utilized discriminator of PatchGAN as discriminator. As for balanced weights, we empirically set α, λ_{Adv}, λ_{Consis}, and λ_{Disen} as 2, 1, 1, and 10.

Table 1. Non-reference evaluations with various dehazing methods on real dataset.

Methods	NIQE↓	User Study↑
DCP [18]	6.71	3.18
ZID [24]	6.63	2.21
DA-Dehazing [42]	5.83	5.73
PSD [9]	5.94	5.12
MSBDN [12]	6.68	4.18
FFA-Net [36]	6.70	4.92
Ours	**5.61**	**6.80**

During training, we randomly crop both synthetic and real images to 256×256 patches. Since directly training from scratch for the whole network is difficult, in which paired ground truth for real haze is unavailable, we firstly pretrain E_{intra} and E_{inter} simultaneously utilizing synthetic data and the corresponding groundtruth for 90 epochs with initial learning rate 1e-4 and Adam optimizer, which serves as a good initialization to ease learning difficulty. The learning rate decrease to zero linearly during pretraining. Then we utilize both synthetic and real data for finetuning the whole network with learning rate 1e-6 for 90 epochs, without learning rate decay. More details can be found in supplementary.

4. Experiments

4.1. Experiment Setting

Dataset. There are various datasets for image dehazing [3–5, 26, 49], but few datasets specially designed for varicolored haze removal. To solve this problem, we establish a large-scale varicolored haze image dataset, which possesses both large-scale synthetic and real-world varicolored haze images. In this paper, we follow [14] and categorize the varicolored haze image into three classes: grayish dense fog, orange smog, and bluish fog. We select 6000 images from RESIDE dataset [26] for haze simulation via Eq. 1. And for each clean background, we generate all three colored haze images with random scattering coefficients, so as to acquire 18000 varicolored haze images for training and 535 images for testing. Moreover, for evaluation on real-world haze images, we collect data from the Internet and existing datasets [10, 26] to acquire 1254 varicolored real-world haze images for training and evaluation. Our code and dataset are publically available at *https://github.com/HuaYuuu/PDI2A-CVPR2022*.

Experimental Settings. We compare PDI2A with (1) prior-based dehazing methods DCP [18], NLD [11]; (2) unsupervised ZID [24]; (3) CycleGAN [48]; (4) semi-supervised Semi-Dehazing [27], DA-Dehazing [42], and PSD [9]; (5) supervised MSBDN [12], FFA-Net [36]. For fair comparisons, we have finetuned all compare methods (if possible) on our dataset with codes provide by the authors.

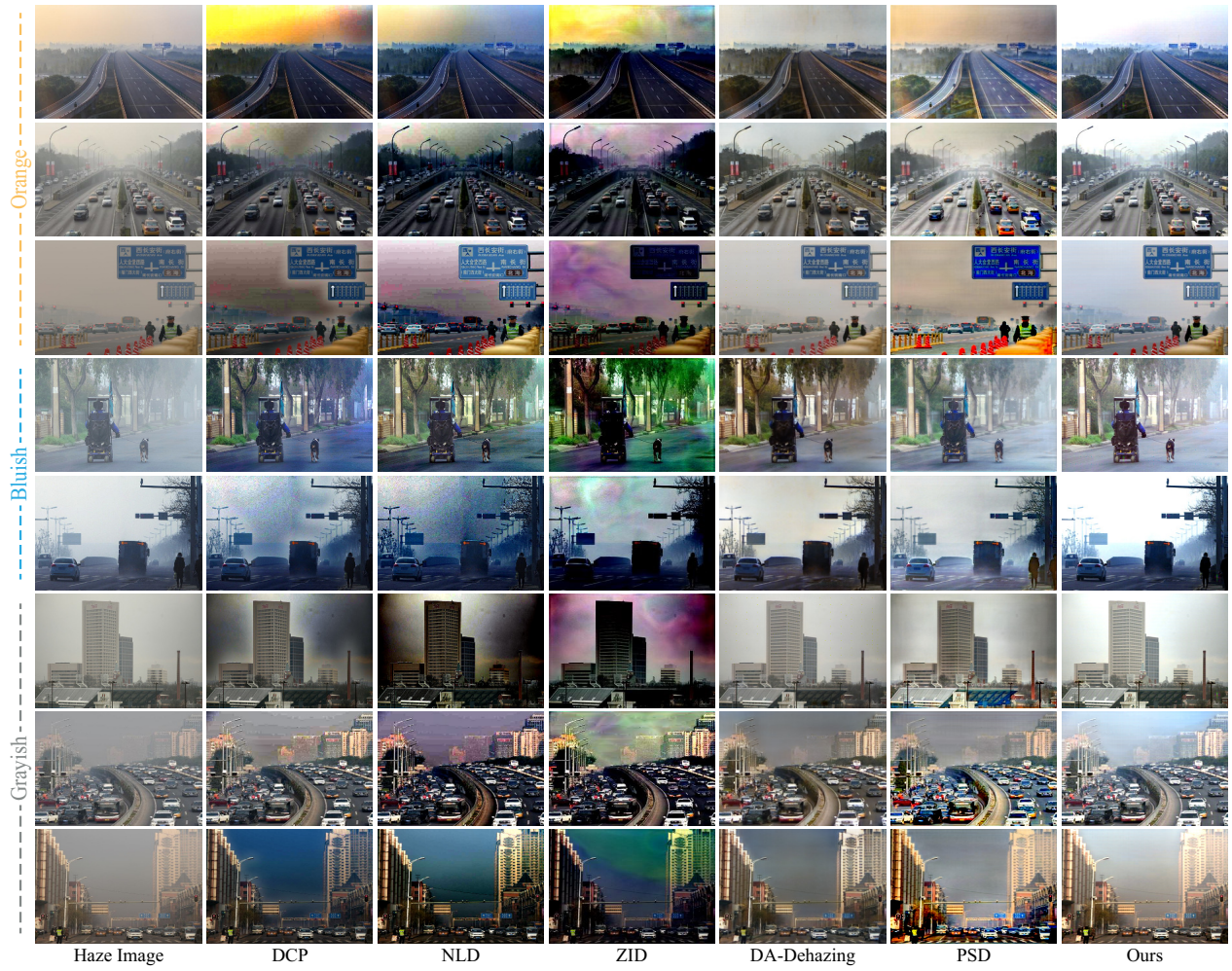

Figure 4. Visualization of dehazing results on real-world varicolored haze images.

4.2. Comparision on Real Data

We evaluate the performance of the proposed PDI2A and other state-of-the-art methods on real varicolored haze data. No-reference metrics such as NIQE and user study are adopted for evaluation. Specifically, for the user study, we randomly choose 50 real varicolored haze images and the corresponding dehazing results produced by each comparison method. Then we invite 20 people to score each result with the range of 1-10, higher score represents better dehazing quality. The results are shown in Tab. 1. We can observe that our results outperform both supervised and other semi-/unsupervised methods by a large margin. As for visual comparison in Fig. 4, DCP [18], NLD [11], and ZID [24] are continually suffer from color distortion due to unsuitable priors for varicolored haze images. DA-Dehazing [42] and PSD [9] can acquire dehazing images with higher contrast, however, results may still encounter slight color distortion and darken scene due to the intra-domain gap. Compared with these methods, PDI2A possesses not only higher contrast but also more satisfying color correction results thanks to the joint intra- and inter-domain adaptation paradigm.

4.3. Comparision on Synthetic Data

We further evaluate the performance of PDI2A and other state-of-the-art semi-/unsupervised methods on synthetic data. The dehazing results visualization are shown in Fig. 5. From second and third columns, we can observe that conventional prior-based methods suffered from severe color distortion and have darkened images, which is because conventional priors [11,18] may not be suitable for varicolored haze. The prior-based method ZID [24] also acquires same results. Moreover, there is still some haze remaining in results produced by Semi-Dehazing [27] and DA-Dehazing [42]. Compared with these methods, PDI2A can acquire images with more favorable visual quality and details, closing to corresponding groundtruth. Moreover, we also perform quantitative comparison of dehaze results in Tab. 2. Our method continues to acquire the highest PSNR/SSIM, demonstrating the superiority of PDI2A for varicolored dehazing.

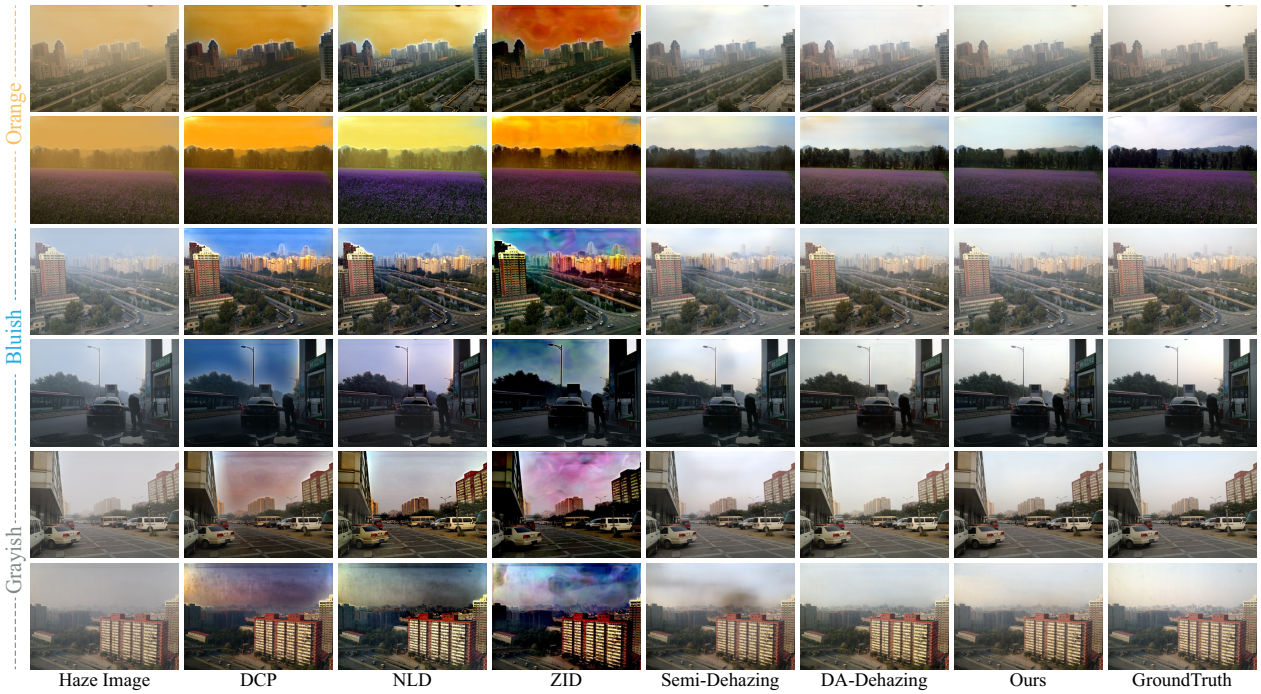

Figure 5. Visualization of dehazing results on synthetic varicolored haze images.

Table 2. Quantitative comparison with various dehazing methods on synthetic dataset.

Methods	PSNR↑	SSIM↑
DCP [18]	15.52	0.8423
NLD [11]	16.32	0.8172
ZID [24]	13.89	0.5804
CycleGAN [48]	24.75	0.9220
Semi-Dehazing [27]	28.70	0.9638
DA-Dehazing [42]	27.88	0.9466
PSD [9]	28.49	0.9477
Ours	**30.13**	**0.9670**

Table 3. The ablation study for components and losses in PDI2A.

Ablations	PSNR	SSIM
w/o intra-DA	29.20	0.9548
w/o inter-DA	29.07	0.9530
w/o L_{Consis}^{Disen}	28.47	0.9498
w/o L_{Consis}^{Recons}	29.94	0.9598
w/o Joint Training	27.95	0.9456
w/o Pretrain	24.95	0.9180
Ours	**30.13**	**0.9670**

4.4. Ablation Study

Effectiveness of intra- and inter-domain adaptation. We investigate the influence of inter-domain adaptation and inter-domain adaptation in Tab. 3. We can observe that both modules serve important roles in quantitative performance. The performance drops in the first two rows demonstrate the benefits brought by intra- and inter-domain adaptation, respectively. Furthermore, we also visualize dehazing result without corresponding module in Fig. 6. We can observe that without color correction in intra-domain adaptation module, the dehazing result could obtain contrast improvement but severe color distortion, which strongly demonstrates importance of intra-domain adaptation when dealing with varicolored haze images. Moreover, without synthetic&real interaction in inter-domain adaptation module, results possess more haze remaining in real scene due to gap between synthetic&real data, which further validate effectiveness of the proposed joint intra- and inter-domain adaptation paradigm.

Effectiveness of consistency losses. We further evaluate the influence of consistency losses, as shown in third and fourth rows. The performance drop clearly demonstrates benefits brought by consistency losses. Moreover, disentangle consistency serves as a strong constraint to disentanglement process, which plays an important role in quantitative results.

Effectiveness of training strategy. Then we evaluate the influence of different training strategies in PDI2A, as shown in the fifth and sixth row in Tab. 3. We can observe the method encounter obvious performance drop without pretraining, denoting the importance of pretraining on synthetic data serving as a good initialization to ease network learning. Moreover, without end-to-end jointly training and learning, the performance has also decreased from 30.13/0.9670 (jointly) to 27.95/0.9456 (separately), showing the benefits of joint intra- and inter-domain adaptation paradigm.

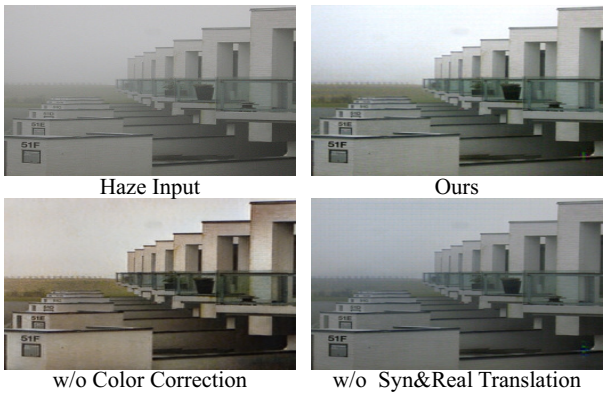

Figure 6. Visualization of dehazing results w/o color correction in intra-domain adaptation and w/o synthetic&real translation in inter-domain adaptation.

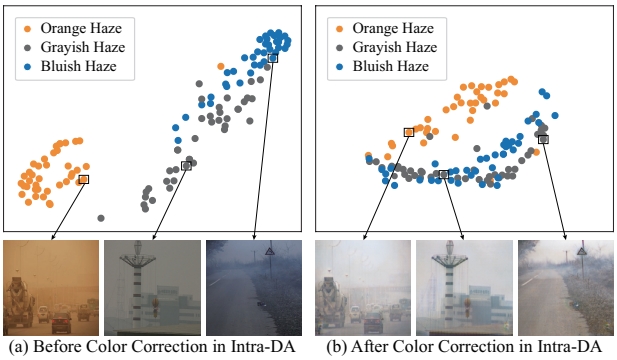

Figure 7. Visualization results of t-SNE on varicolored haze images before and after the color correction in intra-domain adaptation.

4.5. Analysis

Analysis of color correction in feature-level. To study the performance of color correction in inter-domain adaptation, we visualize the t-SNE [44] result of varicolored haze features before and after color correction, as shown in Fig. 7. On one hand, the separation between varicolored haze images demonstrates significant gap among varicolored haze. On the other hand, after color correction, the translated haze images lie on more similar distribution, further demonstrating the importance of intra-domain adaptation.

Analysis of S2R translation. We also investigate the results of synthetic-to-real translation, as shown in Fig. 8. We can observe that the haze images generated from the proposed method are relatively closer to real haze images after translation, revealing the effectiveness of the synthetic and real translation in the inter-domain adaptation process.

4.6. Limitation

The proposed PDI2A mainly bridges the intra- and inter-domain gap simultaneously through physical disentangle and translation, in which atmospheric scattering model [29, 34] plays an important role. However, nighttime haze is far more complicated than daytime haze due to the non-homogeneous illumination and the halo effect. The real-world nighttime haze image and the corresponding results of our PDI2A are shown in Fig. 9. We can observe that our method removes most of haze effects while leaving the severe halo effects behind. Moreover, since the adopted scattering model mainly focuses on no wavelength dependence condition, the proposed PDI2A may face difficulties in wavelength dependence situations such as smog or underwater vision for now. A more comprehensive degradation model would alleviate the problem and we would like to tackle these problems in our future works.

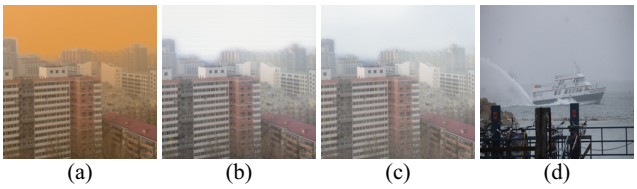

Figure 8. Translated results of synthetic varicolored haze images. (a) Synthetic varicolored haze input, (b) synthetic color balanced image from intra-DA, (c) S2R images, (d) reference real image.

Figure 9. Analysis of the limitation of the proposed PDI2A. Left: real-world nighttime haze images. Right: results produced by PDI2A. Due to huge difference between daytime and nighttime haze images, our method could remove most haze effects while leaving the halo effect behind.

5. Conclusion

In this work, we propose a novel joint intra- and inter-domain adaptation paradigm for varicolored haze removal task, which contains two intra-domain adaptation modules for color correction and one inter-domain adaptation module for synthetic↔real translation. For better translation, we propose to employ a physically disentangle strategy by decomposing the haze image into three components through scattering model, so as to preserve the identity background and ease translation difficulties with only color/haze relevant layer. Extensive experiments on synthetic and real varicolored haze data demonstrate the superiority of our method.

Acknowledgements. This work was supported in part by the National Natural Science Foundation of China under Grant No. 61971460 and No. 62101294, in part by JCJQ Program under Grant 2021-JCJQ-JJ-0060, and in part by the Equipment Pre-Research Foundation under Grant No. 6142113200304.

References

[1] C. Ancuti and C. Ancuti. Single image dehazing by multi-scale fusion. *IEEE Trans. Image Process.*, 22(8):3271–3282, 2013. 2

[2] C. Ancuti, C. Ancuti, and P. Bekaert. Effective single image dehazing by fusion. In *IEEE Int. Conf. Image Process.*, pages 3541–3544, 2010. 2

[3] C. Ancuti, C. Ancuti, and C. De Vleeschouwer. D-hazy: A dataset to evaluate quantitatively dehazing algorithms. In *IEEE Int. Conf. Image Process.*, pages 2226–2230, 2016. 5

[4] C. Ancuti, C. Ancuti, M. Sbert, and R. Timofte. Dense-haze: A benchmark for image dehazing with dense-haze and haze-free images. In *IEEE Int. Conf. Image Process.*, pages 1014–1018, 2019. 5

[5] C. Ancuti, C. Ancuti, and R. Timofte. Ntire 2018 challenge on image dehazing: Methods and results. In *IEEE Conf. Comput. Vis. Pattern Recog. Worksh.*, pages 891–901, 2018. 5

[6] D. Berman, T. Treibitz, and S. Avidan. Air-light estimation using haze-lines. In *IEEE Int. Conf. Comput.*, pages 1–9, 2017. 1

[7] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao. Dehazenet: An end-to-end system for single image haze removal. *IEEE Trans. Image Process.*, 25(11):5187–5198, 2016. 1, 2

[8] D. Chen, M. He, Q. Fan, J. Liao, L. Zhang, D. Hou, L. Yuan, and G. Hua. Gated context aggregation network for image dehazing and deraining. In *Winter Conf. Appl. Comput. Vis.*, pages 1375–1383, 2019. 1, 2

[9] Z. Chen, Y. Wang, Y. Yang, and D. Liu. Psd: Principled synthetic-to-real dehazing guided by physical priors. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 7180–7189, 2021. 1, 2, 5, 6, 7

[10] L. Choi, J. You, and A. Bovik. Referenceless prediction of perceptual fog density and perceptual image defogging. *IEEE Trans. Image Process.*, 24(11):3888–3901, 2015. 5

[11] Shai A. Dana B., Tali t. Non-local image dehazing. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 1674–1682, 2016. 1, 2, 5, 6, 7

[12] H. Dong, J. Pan, L. Xiang, Z. Hu, X. Zhang, F. Wang, and M. Yang. Multi-scale boosted dehazing network with dense feature fusion. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 2157–2167, 2020. 1, 5

[13] J. Dong and J. Pan. Physics-based feature dehazing networks. In *Eur. Conf. Comput. Vis.*, pages 188–204, 2020. 1, 2

[14] A. Dudhane, K. Biradar, P. Patil, P. Hambarde, and S. Murala. Varicolored image de-hazing. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 4564–4573, 2020. 2, 4, 5

[15] R. Fattal. Single image dehazing. *ACM Trans. Graph.*, 27(3):1–9, 2008. 2

[16] R. Fattal. Dehazing using color-lines. *ACM Trans. Graph.*, 34(1):1–14, 2014. 2

[17] A. Golts, D. Freedman, and M. Elad. Unsupervised single image dehazing using dark channel prior loss. *IEEE Trans. Image Process.*, 29:2692–2701, 2019. 1

[18] K. He, J. Sun, and X.u Tang. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(12):2341–2353, 2010. 1, 2, 5, 6, 7

[19] S. Huang, B. Chen, and W. Wang. Visibility restoration of single hazy images captured in real-world weather conditions. *IEEE Trans. Circuit Syst. Video Technol.*, 24(10):1814–1824, 2014. 2

[20] X. Huang and S. Belongie. Arbitrary style transfer in real-time with adaptive instance normalization. In *Int. Conf. Comput. Vis.*, pages 1501–1510, 2017. 4

[21] X. Huang, M. Liu, S. Belongie, and J. Kautz. Multimodal unsupervised image-to-image translation. In *Eur. Conf. Comput. Vis.*, pages 172–189, 2018. 3

[22] H. Lee, H. Tseng, J. Huang, M. Singh, and M. Yang. Diverse image-to-image translation via disentangled representations. In *Eur. Conf. Comput. Vis.*, pages 35–51, 2018. 3, 4

[23] B. Li, Y. Gou, S. Gu, J. Liu, J. Zhou, and X. Peng. You only look yourself: Unsupervised and untrained single image dehazing neural network. *Int. J. Comput. Vis.*, 129(5):1754–1767, 2021. 1

[24] B. Li, Y. Gou, J. Liu, H. Zhu, J. Zhou, and X. Peng. Zero-shot image dehazing. *IEEE Trans. Image Process.*, 29:8457–8466, 2020. 1, 5, 6, 7

[25] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng. Aod-net: All-in-one dehazing network. In *Int. Conf. Comput. Vis.*, pages 4770–4778, 2017. 1, 2

[26] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang. Benchmarking single-image dehazing and beyond. *IEEE Trans. Image Process.*, 28(1):492–505, 2018. 5

[27] L. Li, Y. Dong, W. Ren, J. Pan, C. Gao, N. Sang, and M. Yang. Semi-supervised image dehazing. *IEEE Trans. Image Process.*, 29:2766–2779, 2019. 1, 2, 5, 6, 7

[28] R. Li, J. Pan, Z. Li, and J. Tang. Single image dehazing via conditional generative adversarial network. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 8202–8211, 2018. 1

[29] Y. Li, S. You, M. Brown, and R. Tan. Haze visibility enhancement: A survey and quantitative benchmarking. *Comput. Vis. Image Und.*, 165:1–16, 2017. 1, 8

[30] W. Liu, X. Hou, J. Duan, and G. Qiu. End-to-end single image fog removal using enhanced cycle consistent adversarial networks. *IEEE Trans. Image Process.*, 29:7819–7833, 2020. 1

[31] X. Liu, Y. Ma, Z. Shi, and J. Chen. Griddehazenet: Attention-based multi-scale network for image dehazing. In *Int. Conf. Comput. Vis.*, pages 7314–7323, 2019. 1

[32] Y. Liu, J. Pan, J. Ren, and Z. Su. Learning deep priors for image dehazing. In *Int. Conf. Comput. Vis.*, pages 2492–2500, 2019. 1, 2

[33] Y. Liu, L. Zhu, S. Pei, H. Fu, J. Qin, Q. Zhang, L. Wan, and W. Feng. From synthetic to real: Image dehazing collaborating with unlabeled real data. In *ACM Int. Conf. Multimedia*, pages 50–58, 2021. 1, 2

[34] S. Narasimhan and S. Nayar. Vision and the atmosphere. *Int. J. Comput. Vis.*, 48(3):233–254, 2002. 1, 8

[35] Y. Pei, Y. Huang, Q. Zou, Y. Lu, and S. Wang. Does haze removal help cnn-based image classification? In *Eur. Conf. Comput. Vis.*, pages 682–697, 2018. 1

[36] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Jia. Ffa-net: Feature fusion attention network for single image dehazing. In *AAAI Conf. Artif. Intell.*, volume 34, pages 11908–11915, 2020. 5

[37] Y. Qu, Y. Chen, J. Huang, and Y. Xie. Enhanced pix2pix dehazing network. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 8160–8168, 2019. 1

[38] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley. Color transfer between images. *IEEE Comput. Graph. Appl.*, 21(5):34–41, 2001. 2

[39] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M. Yang. Single image dehazing via multi-scale convolutional neural networks. In *Eur. Conf. Comput. Vis.*, pages 154–169, 2016. 1, 2

[40] W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W. Liu, and M. Yang. Gated fusion network for single image dehazing. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 3253–3261, 2018. 2

[41] C. Sakaridis, D. Dai, and L. Van Gool. Semantic foggy scene understanding with synthetic data. *Int. J. Comput. Vis.*, 126(9):973–992, 2018. 1

[42] Y. Shao, L. Li, W. Ren, C. Gao, and N. Sang. Domain adaptation for image dehazing. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 2808–2817, 2020. 1, 2, 5, 6, 7

[43] R. Tan. Visibility in bad weather from a single image. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 1–8, 2008. 2

[44] L. Van der Maaten and G. Hinton. Visualizing data using t-sne. *J. Mach. Learn. Res.*, 9(11), 2008. 8

[45] H. Wu, Y. Qu, S. Lin, J. Zhou, R. Qiao, Z. Zhang, Y. Xie, and L. Ma. Contrastive learning for compact single image dehazing. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 10551–10560, 2021. 1, 2

[46] X. Yang, Z. Xu, and J. Luo. Towards perceptual image dehazing by physics-based disentanglement and adversarial training. In *AAAI Conf. Artif. Intell.*, volume 32, 2018. 1

[47] S. Zhao, L. Zhang, Y. Shen, and Y. Zhou. Refinednet: A weakly supervised refinement framework for single image dehazing. *IEEE Trans. Image Process.*, 30:3391–3404, 2021. 1

[48] J. Zhu, T. Park, P. Isola, and A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Int. Conf. Comput. Vis.*, pages 2223–2232, 2017. 2, 3, 4, 5, 7

[49] Q. Zhu, J. Mai, and L. Shao. Single image dehazing using color attenuation prior. In *Brit. Mach. Vis. Conf.*, 2014. 5

[50] Q. Zhu, J. Mai, and L. Shao. A fast single image haze removal algorithm using color attenuation prior. *IEEE Trans. Image Process.*, 24(11):3522–3533, 2015. 1