

Source-Free Domain Adaptation for Real-World Image Dehazing

Hu Yu*
University of Science and Technology
of China
yuhu520@mail.ustc.edu.cn

Qi Zhu
University of Science and Technology
of China
zqcrafs@mail.ustc.edu.cn

Jie Huang*
University of Science and Technology
of China
hj0117@mail.ustc.edu.cn

Man Zhou
University of Science and Technology
of China
manman@mail.ustc.edu.cn

Yajing Liu
JD Logistics
liuyajing25@jd.com

Feng Zhao†
University of Science and Technology
of China
fzhao956@ustc.edu.cn

ABSTRACT

Deep learning-based source dehazing methods trained on synthetic datasets have achieved remarkable performance but suffer from dramatic performance degradation on real hazy images due to domain shift. Although certain Domain Adaptation (DA) dehazing methods have been presented, they inevitably require access to the source dataset to reduce the gap between the source synthetic and target real domains. To address these issues, we present a novel Source-Free Unsupervised Domain Adaptation (SFUDA) image dehazing paradigm, in which only a well-trained source model and an unlabeled target real hazy dataset are available. Specifically, we devise the Domain Representation Normalization (DRN) module to make the representation of real hazy domain features match that of the synthetic domain to bridge the gaps. With our plug-and-play DRN module, unlabeled real hazy images can adapt existing well-trained source networks. Besides, the unsupervised losses are applied to guide the learning of the DRN module, which consists of frequency losses and physical prior losses. Frequency losses provide structure and style constraints, while the prior loss explores the inherent statistic property of haze-free images. Equipped with our DRN module and unsupervised loss, existing source dehazing models are able to dehaze unlabeled real hazy images. Extensive experiments on multiple baselines demonstrate the validity and superiority of our method visually and quantitatively.

CCS CONCEPTS

• Computing methodologies → Scene understanding.

KEYWORDS

Single image dehazing, source-free, domain adaptation, domain knowledge disentangling

*Equal contribution.

†Corresponding author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MM '22, October 10–14, 2022, Lisboa, Portugal
© 2022 Association for Computing Machinery.
ACM ISBN 978-1-4503-9203-7/22/10...\$15.00
<https://doi.org/10.1145/3503161.3548410>

ACM Reference Format:

Hu Yu, Jie Huang, Yajing Liu, Qi Zhu, Man Zhou, and Feng Zhao. 2022. Source-Free Domain Adaptation for Real-World Image Dehazing. In *Proceedings of the 30th ACM International Conference on Multimedia (MM '22), October 10–14, 2022, Lisboa, Portugal*. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3503161.3548410>

1 INTRODUCTION

Haze is a common atmospheric phenomenon. Images captured in hazy environments usually suffer from noticeable visual quality degradation in object appearance and contrast, resulting in accuracy decreasing for subsequent visual analysis. Thus, image dehazing has been a focus of research in the computational photography and vision communities throughout the last decades.

As recognized, the hazing process can be represented by the physical scattering model [34], which is usually formulated as

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

where $I(x)$ and $J(x)$ denote the hazy image and the clean image respectively, A is the global atmospheric light, and $t(x)$ is the transmission map.

However, estimating the clean image from a single hazy input is an ill-posed and challenging problem. Given a hazy image $I(x)$, conventional prior-based dehazing algorithms attempt to estimate $t(x)$ and A by constraining the solution space using a variety of sharp image priors [3, 12–14] and then restore the image via the scattering model. However, these hand-crafted image priors are based on specific observations, which may not be reliable for estimating the transmission map in the physical scattering model.

Numerous convolutional neural networks (CNNs)-based systems have been developed to estimate transmission maps [4, 43, 57] or yield clean images directly [9, 23, 27, 30, 41, 42, 44, 52], which achieve superior image dehazing performance over classic prior-based algorithms. However, these approaches require large quantities of paired hazy/clean images to train in a supervised learning manner. In general, due to the impracticality of acquiring large amounts of hazy-clean pairs in the real world, most dehazing models are trained on hazy synthetic datasets. In this paper, we denote these synthetic hazy dataset trained models as **source models**. However, source models often suffer from degraded performance on real-world hazy images due to the domain gap.

Recently, the above domain shift problem has drawn the attention of the image dehazing community. Existing Domain-Adaptation (DA) dehazing methods [6, 21, 22, 25, 33, 46, 56] achieve remarkable

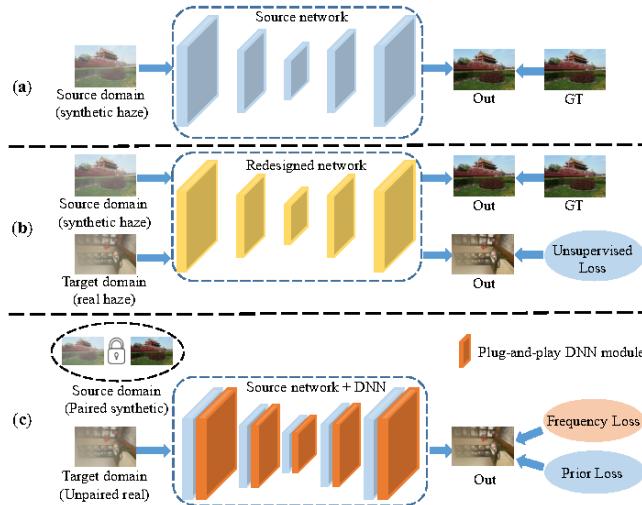


Figure 1: (a) Source dehazing models that are trained on synthetic data (e.g. MSBDN [9]). (b) Existing source-driven DA dehazing methods. These methods design customized architecture and need both source synthetic and target real hazy image during domain adaptation. (c) Our Source-Free Unsupervised Domain Adaptation image dehazing paradigm. Our method works with only target unpaired real hazy images available and can direct utilize well trained source models with frozen parameters in a plug-and-play manner.

performance, but they have two common drawbacks. First, they inevitably require full access to source synthetic hazy datasets to reduce the gap between the source synthetic and target real domains during model adaptation, which is cumbersome in practice due to storage, privacy, and transmission constraints. Second, most of these methods need to retrain source models with modification or use customized architecture for domain adaptation, neglecting to direct utilize numerous source models, which reduces the generality of their methods. In this paper, we denote these DA dehazing methods as **source-driven models**.

To address these problems, we present a novel Source-Free Unsupervised Domain Adaptation (SFUDA) image dehazing paradigm, where only a well-trained source model and an unlabeled target real hazy dataset are available during model adaptation. The overview of our method is shown in Fig. 1. Since the representations of features across synthetic and real domains vary significantly, our SFUDA achieves domain adaptation from the domain representation perspective. Specifically, the DRN module is devised for making the representation of real hazy image's features match that of the synthetic domain to adapt the source network. To implement this, the DRN module disentangles the features of real hazy images into domain-invariant and domain-variant parts. The domain-invariant part is obtained by Instance Normalization (IN), while the domain-variant part is implicitly guided by the feature statistics to make the features adapt to the source model. With IN equipped in the feature space, it can normalize feature statistics for style normalization [18]. The proposed DRN module can be directly applied to existing dehazing models in a plug-and-play fashion.

Additionally, the unsupervised frequency losses and physical prior losses are applied to regularize the dehazing processing of

unlabeled real hazy images. Specifically, we explore the frequency property in our unsupervised setting. It is worth noting that although source network fails to remove the haze of real images, it well reconstructs the structural information. Therefore, we introduce the phase structure loss between the intermediate stages and output of the source network and the student network for structural consistency. Besides, we find that the original real hazy image has an enhanced illumination and color contrast after being processed by Contrast Limited Adaptive Histogram Equalization (CLAHE). With this property, we introduce the amplitude style loss to regularize the enhanced illumination contrast between the output of CLAHE and the output of the student network. Moreover, we employ the classical Dark Channel Prior (DCP) and Color Attenuation Prior (CAP) to form prior loss to exploit the inherent statistic property of haze-free images.

Existing source models that are equipped with our DRN module and unsupervised losses gain generalization power and can dehaze unlabeled real hazy images. In conclusion, the main contributions of this work are summarized as follows:

- We propose a novel Source-Free Unsupervised Domain Adaptation (SFUDA) image dehazing paradigm. To the best of our knowledge, this is the first attempt to address the problem of source-free UDA for image dehazing. Our method can be directly applied to existing dehazing models in a plug-and-play fashion, which is more general in practical.
- we propose the Domain Representation Normalization (DRN) module to regularize representations of features in the real hazy domain to adapt to the frozen source network.
- To train the whole framework in an unsupervised manner, we leverage the frequency property and physical priors to constrain the adaptation and generate more natural images.
- Extensive experiments on multiple baselines demonstrate the validity of our method visually and quantitatively. In particular, our method outperforms the state-of-the-art source-driven UDA approaches under a source-free setting.

2 RELATED WORK

2.1 Image dehazing

Recent years have witnessed significant advances in image processing [39, 51, 54], including single image dehazing. Existing image dehazing methods can be roughly categorized into physical-based methods and deep learning-based methods.

Physical-based. Physical-based methods depend on the physical model [34] and the handcraft priors from empirical observation, such as dark channel prior [14], color line prior [13], and sparse gradient prior [5]. However, the density of haze can be affected by various factors, which makes the haze formation at individual spatial locations space-variant. Therefore, the haze usually cannot be accurately characterized by merely a single transmission map.

Deep learning-based. Different from the physical-based methods, deep learning-based methods employ convolution neural networks and large-scale datasets to learn the image prior [4, 23, 29, 31, 43, 57] or directly learn hazy-to-clear image translation [8–10, 16, 30, 41, 42, 44, 50, 52, 55, 59]. For example, MSBDN [9] proposes a boosted decoder to progressively restore the haze-free images. Existing methods have shown outstanding performances on

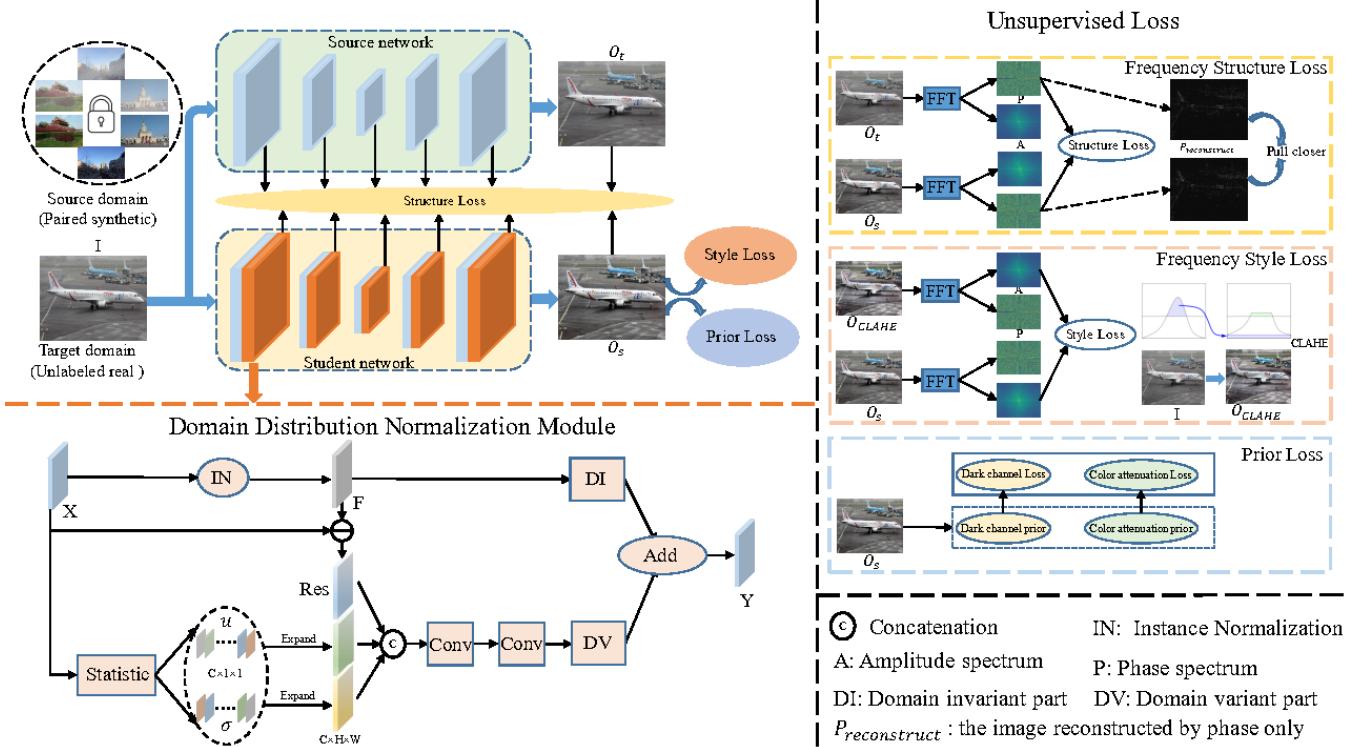


Figure 2: Overview of the proposed framework. The source network can be any synthetic data pretrained model (e.g., MSBDN), while the student network is built by inserting our DRN module into the intermediate stages of the source model. Note that the parameters in the teacher network and student network’s source model part are frozen.

image dehazing. However, they train on paired synthetic data and generalize poorly on real-world data.

2.2 Domain adaptation dehazing

Recently, some domain adaptation dehazing methods [6, 21, 22, 25, 33, 46, 47, 56] have been proposed to tackle domain shift between synthetic and real domains. For example, Shao et al. [46] developed a domain adaptation paradigm, which consists of a image translation module and two dehazing modules and the image translation module is used for data augmentation. Chen et al. [6] modified and retrained an existing dehazing model with physical priors and fine-tune it on both synthetic and real hazy images.

However, these existing techniques require the full access to source synthetic hazy datasets during model adaptation, which limits their practical application, due to the non-availability of source datasets in some cases.

2.3 Source-Free Unsupervised Domain Adaptation

Since labeled source data may not be available in some real-world scenarios due to data privacy issues, Source-Free Unsupervised Domain Adaptation aims to explore how to improve performance of an existing source model on the target domain with only unlabeled target data available. Recently, this new domain adaptation paradigm has been applied to different tasks [1, 7, 17, 26, 32] for its high practical value and easy-to-use property. Different from

these techniques, we first attempt to introduce SFUDA to image dehazing tasks, and innovatively address this problem from the domain representation perspective. Our solution is general and can be directly applied to existing source dehazing models.

3 METHOD

3.1 Method Overview

Source dehazing models are trained on synthetic data, thus obtaining the dehazing knowledge for the representation of synthetic domain. With this property in mind, we design the DRN module from the domain representation perspective to adapt to the source model. Besides, we design the unsupervised loss to constrain the learning process. Our method makes the first attempt to directly leverage the learned dehazing knowledge of existing source models, providing a general solution for domain adaptation and works in a source-free and plug-and-play manner.

As shown in Fig. 2, our domain adaptation framework works in a teacher-student mode. The teacher network can be any existing source dehazing models (e.g. FFA-Net [41], MSBDN [9], we choose MSBDN as the source model by default), while the student network consists of our DRN module and the source model. DRN module is expected to make the representation of real data match the synthetic domain to adapt the frozen source network. Besides, although the source network performs poorly on real hazy images, it owns the ability to preserve the structure of images. Thus, we froze the

parameter of the source network during training so that the structure information of the source network can provide supervision for the student network. The parameter of the student network has two parts: the DRN module part and the source network part. Only the parameter of the DRN module is updated during back-propagation to keep the dehazing ability of the source model on synthetic domain representation.

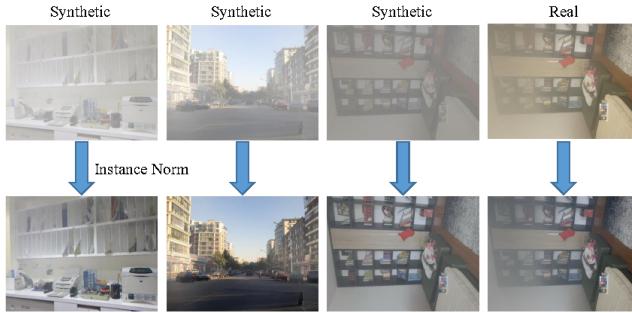


Figure 3: Examples of both real and synthetic hazy images after Instance Normalization operation.

3.2 Domain Representation Normalization Module

Existing normalization-based methods [11, 19] regularize the distribution by normalize it to the standard normal distribution and then modulate it to another distribution with learned statistics. After such $\sigma_1 \frac{X-\mu}{\sigma} + \mu_1$ operation, the distribution change from original $N(\mu, \sigma^2)$ to learned $N(\mu_1, \sigma_1^2)$, where $N(\mu, \sigma^2)$ denotes normal distribution with mean μ and standard deviation σ . However, although such a strategy works well for high-level tasks, normalization will inevitably induce the loss of the image discriminative features [20, 38] for image reconstruction.

Differently, for input real hazy domain feature $X(x) \in \mathbb{R}^{C \times H \times W}$, where C , H and W represent channel dimension, height, and width, respectively, our DRN module disentangles it into domain invariant knowledge DI and domain variant knowledge DV . We introduce Instance normalization (IN) to get DI . IN performs style normalization by normalizing feature statistics, which have been found to carry the style information of an image [18].

We show some hazy images after IN operation in Fig. 3. The synthetic and real hazy images of the same scene are pulled closer in style after IN, which demonstrates that IN property of excavating domain-invariant features fits image dehazing task.

For the domain invariant part, $X(x)$ is first processed by IN to get $F(x)$, formulated by:

$$F(x) = \text{InstanceNorm}(X) = \frac{X - \mu}{\sigma}, \quad (2)$$

where, $\mu \in \mathbb{R}^{C \times 1 \times 1}$ and $\sigma \in \mathbb{R}^{C \times 1 \times 1}$ denote mean and standard deviation of $X(x)$ along the channel dimension. $F(x)$ represents domain-invariant features of standard normal distribution.

For the domain variant part, the residual between input $X(x)$ and domain invariant part $F(x)$ are denoted as $Res(x)$, which can be regarded as domain variant part. Besides, we also obtain the

mean and standard deviation of $X(x)$.

$$\mu = \frac{1}{m} \sum_{k \in S_i} x_k, \quad \sigma = \sqrt{\frac{1}{m} \sum_{k \in S_i} (x_k - \mu_i)^2 + \epsilon} \quad (3)$$

with ϵ be a small constant for numerical stability. Unlike previous methods, we implicitly incorporate μ and σ . Concretely, we first expand μ and σ along the spatial dimension to get the same size as $Res(x)$. Afterwards, the expanded μ , σ and $Res(x)$ are concatenated along the channel dimension to implicitly guide the learning of $Res(x)$. The final domain variant knowledge DV can be computed as:

$$DV = \text{Conv}(\text{Cat}(\text{Res}, \mu, \sigma)). \quad (4)$$

Finally, we integrate domain invariant knowledge DI with domain variant knowledge DV to get the final output $Y(x)$ of the DRN module. Our DRN module keeps domain invariant knowledge DI and modifies the domain variant knowledge along with the loss constraints in Sec. 3.3, which make the representation of output $Y(x)$ match the synthetic representation. To better testify this, we show the feature visualization in Fig. 4. It is evident that real hazy image features in our student network are more similar to the synthetic domain representation than real hazy image features in the source network.

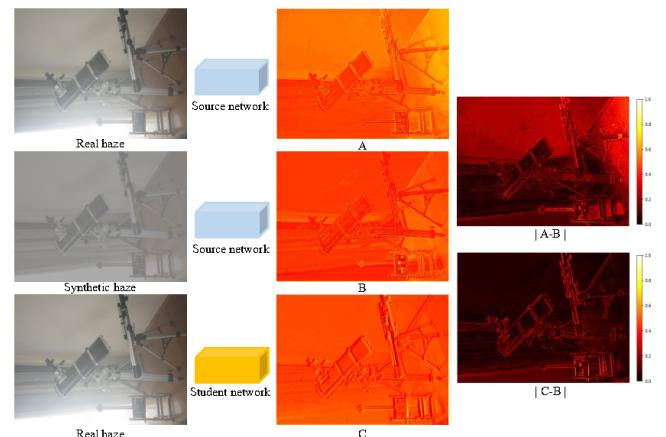


Figure 4: Feature visualization on the effectiveness of our DRN module. The second column is intermediate features.

3.3 Training Losses

Our framework works in an unsupervised way. Therefore, it is critical to design proper loss function as supervision to drive the DRN module learning. Specifically, We explore the frequency property in our unsupervised setting and innovatively introduce two frequency losses for keeping domain-invariant structure and modulating domain-variant style. Besides, we select two effective and well-grounded physical priors to provide us with the prior knowledge of real images.

3.3.1 Frequency domain Losses. As is known, the illumination contrast of an image is represented by the amplitude spectrum, while the structure information is represented by the phase spectrum [37, 48]. Based on this theory, we further conduct an amplitude and phase exchange experiment to validate its applicability in our setting. As shown in Fig. 5, we exchange the amplitude and phase

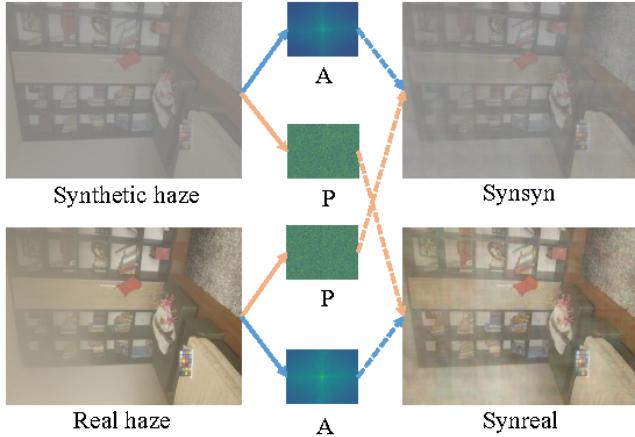


Figure 5: Visualization on the relationship between the haze degradation and the characteristics of amplitude spectrum and phase spectrum in frequency domain.

of synthetic and real hazy images. The results demonstrate that synthetic and real hazy images of the same scene have approximately the same phase (structure) spectrum, while the style and haze degradation mainly manifests in the amplitude spectrum.

Phase (Structure) Loss. The source dehazing network can not generalize well to real hazy images, but they can still retain satisfactory structural information, which is an important dehazing prior to unsupervised learning. With this prior knowledge in mind, we design a Phase (Structure) Loss.

Concretely, we explore the information contained in the phase spectrum of both intermediate features and output images from teacher to student to provide structural constraints. The outputs of the teacher and student network are denoted as O_t and O_s . We first transform the output image to the frequency domain by Fourier Transform.

$$F(x)(u, v) = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} x(h, w) e^{-j2\pi(\frac{h}{H}u + \frac{w}{W}v)}. \quad (5)$$

The frequency-domain feature $F(x)$ is denoted as $F(x) = R(x) + jI(x)$, where $R(x)$ and $I(x)$ represent the real and imaginary part of $F(x)$. Then the real and imaginary parts are converted to amplitude and phase spectra, which can be formulated as:

$$\begin{aligned} A(x)(u, v) &= [R^2(x)(u, v) + I^2(x)(u, v)]^{1/2}, \\ P(x)(u, v) &= \arctan \left[\frac{I(x)(u, v)}{R(x)(u, v)} \right], \end{aligned} \quad (6)$$

where $A(x)$ is the amplitude spectrum, $P(x)$ is the phase spectrum. The phase spectra of teacher and student network outputs are denoted as $P_t(x)$ and $P_s(x)$. Then the Phase (Structure) Loss are formulated as:

$$L_{Pha} = \frac{2}{UV} \sum_{u=0}^{U/2-1} \sum_{v=0}^{V-1} \| |P_t|_{u, v} - |P_s|_{u, v} \|_1. \quad (7)$$

Note, in our implementation, the summation for u is only performed up to $U/2 - 1$, since 50% of all frequency components are redundant. Moreover, this phase loss is also applied to intermediate features in the same way.

Amplitude (Style) Loss. Contrast Limited Adaptive Histogram Equalization (CLAHE) [40, 45] is a classical digital image processing

technology for improving image contrast and it is also applied to image dehazing. However, it is inadvisable to directly apply CLAHE-enhanced images as supervision since CLAHE inevitably introduces bottom noise and brings inherent flaws into our network. To address this problem, we leverage the mentioned frequency property that the illumination contrast of an image is represented by the amplitude spectrum and design the Amplitude (Style) Loss to utilize CLAHE-enhanced illumination and contrast.

Specifically, the amplitude spectrum of student network output $O_s(x)$ is denoted as $A_s(x)$ and the amplitude spectrum of CLAHE-enhanced real hazy image is denoted as $A_{CLAHE}(x)$. Then the Amplitude (Style) Loss is formulated as:

$$L_{Amp} = \frac{2}{UV} \sum_{u=0}^{U/2-1} \sum_{v=0}^{V-1} \| |A_s|_{u, v} - |A_{CLAHE}|_{u, v} \|_1. \quad (8)$$

3.3.2 Physical Prior Losses. Various handcraft priors, such as dark channel prior [14], color line prior [13], color attenuation prior [60], sparse gradient prior [5], maximum reflectance prior [58], and non-local prior [3], are drawn from empirical observation and reflect inherent statistic property of haze-free images. From these priors, we select two effective and well-grounded ones to provide us with the prior knowledge of real images. Additionally, unlike [6, 33], we employ these prior losses with only output images available.

Dark Channel Prior (DCP) Loss. Dark Channel Prior (DCP) [14] is the most famous and effective prior for image dehazing. DCP is a statistical property of outdoor haze-free images: most patches in these images should contain pixels that are dark in at least one color channel. We formulate DCP as a loss function as in [22].

$$L_{DCP} = \left\| \min_{c \in \{r, g, b\}} (F^c(y)) \right\|_p, \quad (9)$$

where $F^c(\cdot)$ is the c -th color channel of y , and y is a local patch of the student network output $O_s(x)$. With this dark channel loss, our student network incorporates the statistical properties from the recovered clean images, avoiding an explicit ground truth on the recovered image.

Color attenuation prior (CAP) Loss. Color attenuation prior (CAP) [60, 61] can be explained as, in a haze-free region, the difference between the brightness and the saturation is close to zero. In concrete implementation, the difference between the value and saturation in the predicted $O_s(x)$ should be minimized as small as possible. We formulate CAP as a loss function as in [21].

$$L_{CAP} = \| V(O_s(x)) - S(O_s(x)) \|_p, \quad (10)$$

where $V(O_s(x))$ and $S(O_s(x))$ respectively denotes the brightness and saturation of $O_s(x)$.

Total Loss. The final loss function used in our method can be defined as:

$$L = \lambda_p L_{Pha} + \lambda_a L_{Amp} + \lambda_d L_{DCP} + \lambda_c L_{CAP}, \quad (11)$$

where λ_p , λ_a , λ_d and λ_c are weight factors.

4 EXPERIMENTS

4.1 Experiment Setup

Datasets. We select Unannotated Real Hazy Images (URHI) from RESIDE dataset [24] for domain adaptation training. For evaluating the effectiveness of our method, we test on two real-world datasets,

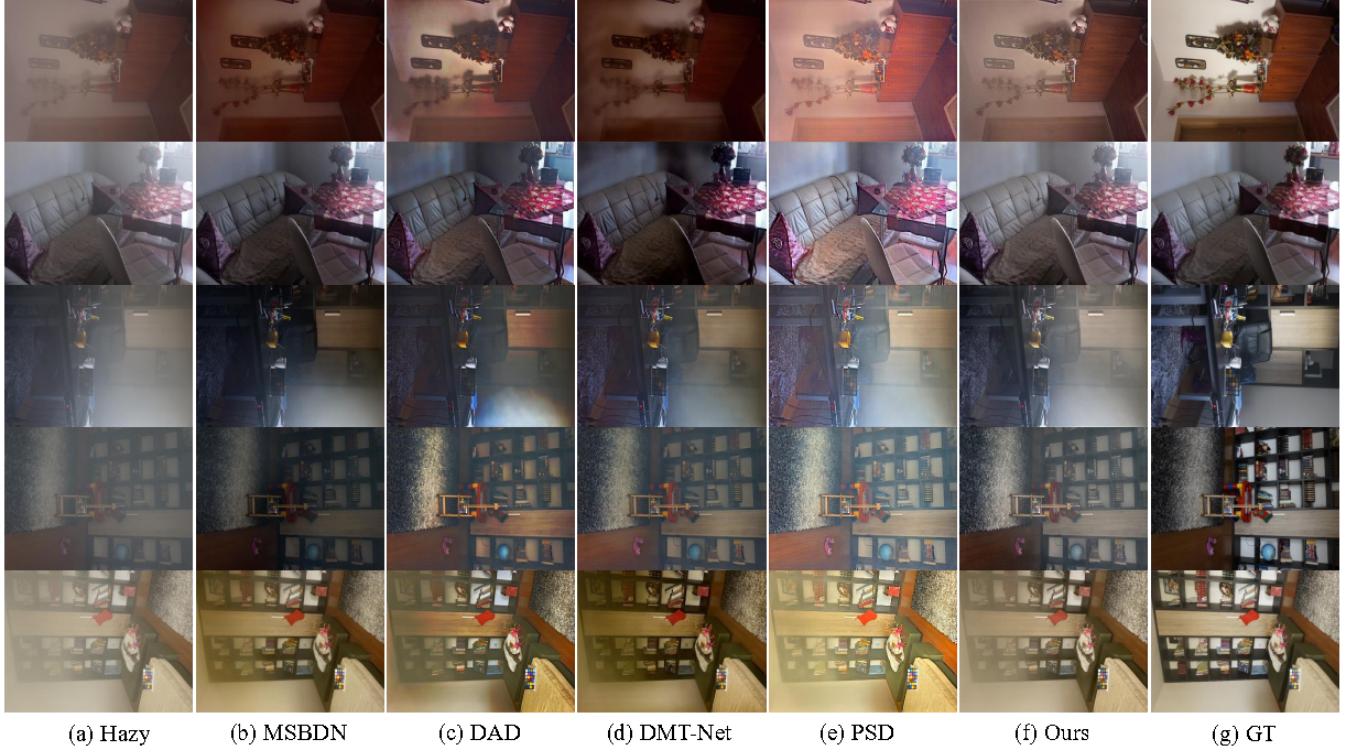


Figure 6: Visual comparison of source baseline and DA dehazing results on labeled real hazy dataset I-Haze.

RTTS [24] and I-Haze [2]. RTTS is a subset of RESIDE dataset [24], consisting of 4322 real unlabeled outdoor hazy images. I-Haze contains 30 paired real indoor hazy images.

Metrics. We employ two widely used metrics, the Peak Signal to Noise Ratio (PSNR) and the Structural Similarity Index (SSIM), to quantitatively assess the results on the labeled I-Haze dataset. Besides, we utilize two well-known no-reference image quality assessment indicators: BRISQUE [35] and NIQE [36] to assess the results on the unlabeled RTTS dataset.

Implementation Details. We use ADAM as the optimizer with $\beta_1 = 0.9$, and $\beta_2 = 0.999$, and the initial learning rate is set to 1×10^{-4} . The learning rate is adjusted by the cosine annealing strategy [15]. The training epoch, batch and patch sizes are set to 10, 6, and 256×256 , respectively. The trade-off weights in loss function are set to $\lambda_p = 1$, $\lambda_a = 1$, $\lambda_d = 10^{-3}$, and $\lambda_c = 10^{-3}$.

Table 1: Quantitative comparisons with state-of-the-art (SOTA) DA methods on two real-world dehazing datasets. ↓ denotes lower is better, while ↑ means higher is better.

	Method	I-HAZE [2]		RTTS [24]	
		PSNR↑	SSIM↑	BRISQUE↓	NIQE↓
	Hazy	-	-	36.703	5.209
Baseline	MSBDN [9]	16.623	0.780	32.575	4.865
DA	DAD [46]	12.076	0.750	34.423	5.367
DA	DMT [33]	12.902	0.521	31.594	4.963
DA	PSD [6]	14.575	0.781	28.011	4.494
DA	Ours	17.602	0.802	27.330	4.326

4.2 Comparison with State-of-the-art Methods

We compare the performance of our SFUDA with three SOTA DA dehazing methods: DAD [46], PSD [6], and DMT-Net [33]. Amounts of experiments are conducted on two real-world hazy datasets.

4.2.1 Comparison on labeled I-Haze Dataset. **Visual Quality.** As shown in Fig. 6, Compared with the ground truths, it is evident that the results of MSBDN, DAD, and DMT not only fail to remove the dense haze but also suffer from color shift. Moreover, the severe color distortion problem of PSD is clearly exposed under ground-truth haze-free images. Compared with all these methods, our method generates the highest-fidelity dehazed results.

Quantitatively results. Table 1 compares the quantitative results of different methods on the I-Haze dataset, which indicates our source-free method achieves the best performance with 17.602dB PSNR and 0.802 SSIM. What's more, different from outdoor datasets URHI and RTTS, I-Haze is an indoor dataset and its representation is far away from that of the training dataset URHI than the RTTS dataset, which is shown in the supplementary material. Consequently, our best performance on the I-Haze dataset strongly proves the generalization ability of our method.

4.2.2 Comparison on unlabeled RTTS Datasets. **Visual Quality.** As shown in Fig. 7, the source baseline model MSBDN remains haze residual, especially in distant areas. The dehazing results of DAD and DMT are dark in some regions. PSD tends to produce visually satisfactory images, but it over amplifies color contrast by direct use of CLAHE prior, thus seems to be unnatural. In contrast,

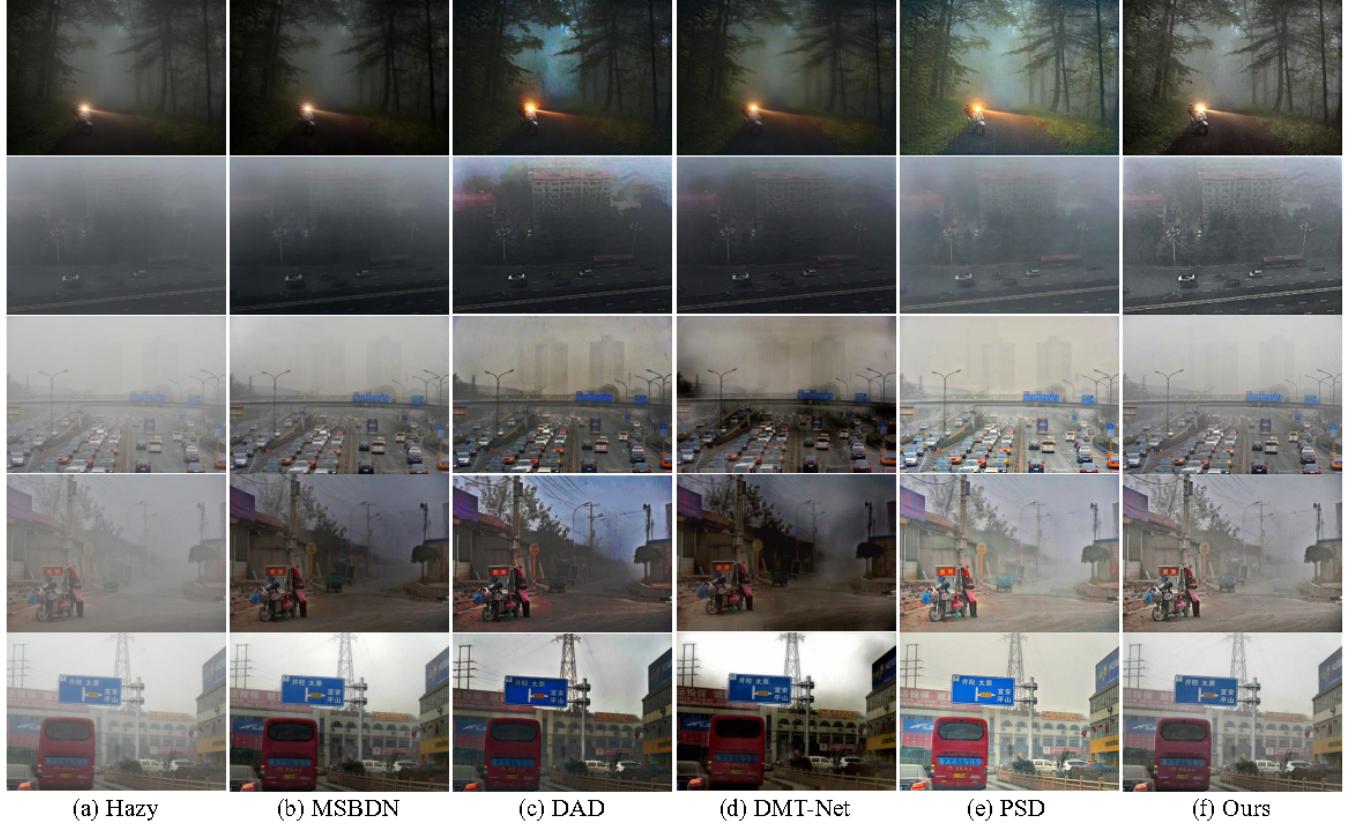


Figure 7: Visual comparison of DA dehazing results on unlabeled real hazy dataset RTTS.

our SFUDA generates high-quality haze-free images with more natural color, clearer architecture, and finer details.

No-Reference Image Quality Assessment. For quantitative comparison of the unlabeled dataset, we employ two well-known no-reference image quality assessment indicators: BRISQUE [35] and NIQE [36]. As shown in table 1, our method largely outperforms baseline MSBDN. Besides, it also surpasses the source-driven DA methods DAD, DMT, and PSD.

Task-Driven Evaluation. The performance of high-level computer vision tasks such as object detection and scene understanding is considerably influenced by input images captured in hazy scenes [28, 49, 53]. Thus, image dehazing task can be used as a preprocessing step for these high-level tasks. Inversely, the performance of these downstream tasks can also be used to evaluate the effectiveness of image dehazing algorithms. To this end, we testify our method and other SOTA dehazing methods on the RTTS dataset with object detection task. We show the detection result of a representative image sampled from RTTS in Figure 8. It is evident that the detection performance of our method surpasses the baseline MSBDN and DA dehazing methods DAD and DMT-Net. Besides, it achieves comparable results compared to PSD.

4.3 Comparison with fine-tuned source models

In order to compare more fairly with existing source baseline methods and further prove the superiority of our method, we compare with two state-of-the-art source baselines methods: FFA-Net [41]

Table 2: Quantitative comparisons with baseline methods and their vanilla direct fine-tuning on two real-world dehazing datasets. “ft” represents fine-tuned baseline.

Method	RTTS [24]		I-HAZE [2]	
	BRISQUE↓	NIQE↓	PSNR↑	SSIM↑
FFA-Net	59.145	8.212	10.105	0.552
FFA-Net (ft)	42.145	7.393	12.151	0.638
Ours (FFA-Net)	29.311	6.191	13.404	0.718
MSBDN	32.575	4.865	16.623	0.780
MSBDN (ft)	30.210	4.659	17.180	0.781
Ours (MSBDN)	27.330	4.326	17.602	0.802

and MSBDN [9] and their fine-tuned ones. Note that these source models are fine-tuned in the same setting as our method for fair comparison. The quantitative comparison results are presented in Table 2. It is obvious that applying our method to these baseline methods is much more effective than vanilla direct fine-tuning. Not that our method improves performance with negligible additional parameters (MSBDN 31.35M, Ours(MSBDN) 31.93M).

Besides, we present the visual comparison in Fig. 9, vanilla direct fine-tuning ones still fails to remove the haze and suffer from color shift problem. In contrast, our method on these baselines generates natural and visually desirable results.

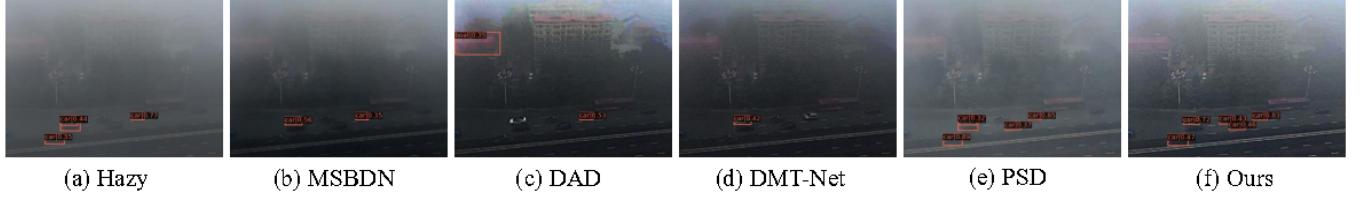


Figure 8: Detection results of SOTA DA methods on a representative image sampled from RTTS dataset.

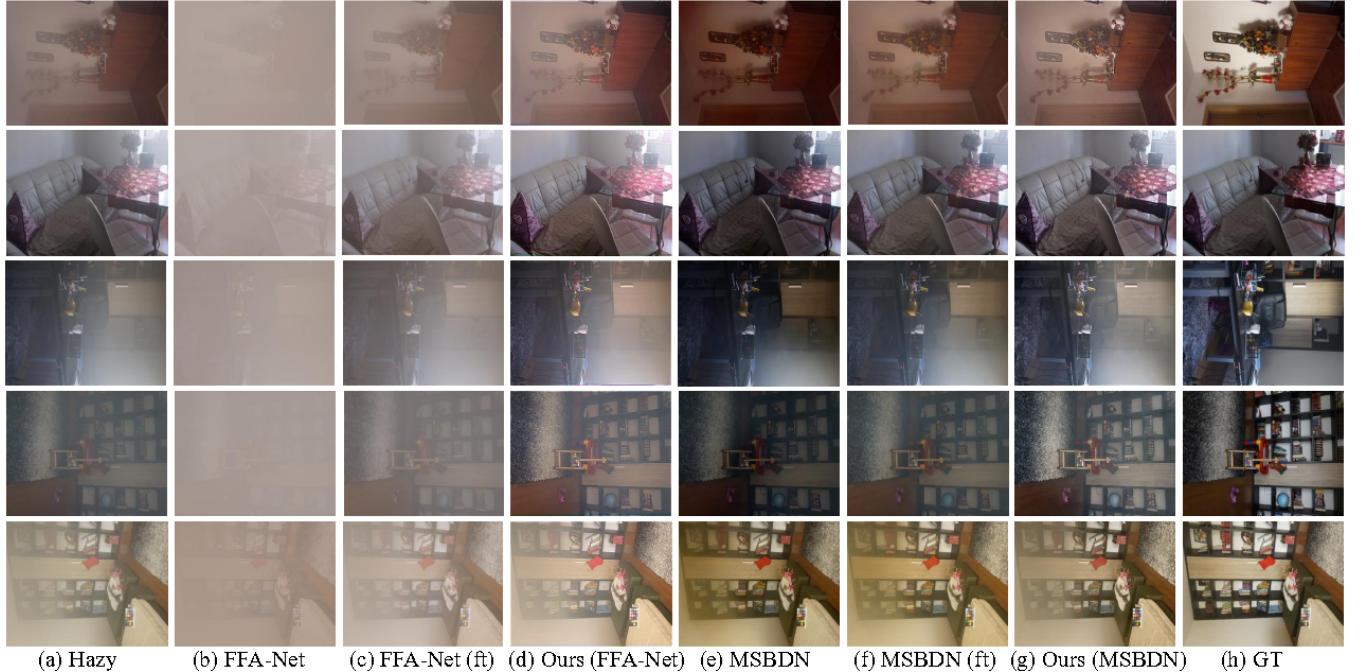


Figure 9: Visual comparison with baselines and their fine-tuned (ft) models on labeled real hazy dataset I-Haze.

Table 3: Ablation study of our method on I-Haze dataset.

Label	DRN	Str	Sty	DCP	CAP	PSNR (dB)	SSIM
a	✗	✓	✓	✓	✓	17.180	0.781
b	✓	✗	✓	✓	✓	17.124	0.781
c	✓	✓	✗	✓	✓	16.752	0.780
d	✓	✓	✓	✗	✓	17.548	0.794
e	✓	✓	✓	✓	✗	17.564	0.796
f	✓	✓	✓	✓	✓	17.602	0.802

4.4 Ablation Studies

In this section, we perform ablation study experiments to evaluate the effectiveness of major components of our method. Specifically, we denote our frequency structure loss, frequency style loss, physical DCP loss, and physical CAP loss as "Str", "Sty", "DCP", and "CAP", respectively. Different models are denoted as follows: (a) without DRN module; (b) without structure loss; (c) without style loss; (d) without DCP loss; (e) without CAP loss; (f) our method. Table 3 summarizes the PSNR and SSIM results of our method under the above settings. It is obvious that our designed DRN module, Style loss, and Structure loss are of greater importance. We further verify the rationality of the DRN module in the supplementary material.

5 CONCLUSION

In this paper, we present a novel Source-Free Unsupervised Domain Adaptation image dehazing paradigm, in which only a well-trained source model and an unlabeled target real hazy dataset are available. Our SFUDA achieves domain adaptation from the domain representation perspective and can be directly applied to existing dehazing models in a plug-and-play fashion. Specifically, we devise a Domain Representation Normalization(DRN) module to make the representation of real hazy data match the synthetic hazy domain. Besides, we leverage the frequency property and physical prior knowledge to form unsupervised losses. Extensive experiments validate that the proposed method achieves SOTA domain adaptation performances in a source-free and easy-to-use way.

ACKNOWLEDGMENTS

This work was supported by the JKW Research Funds under Grant 20-163-14-LZ-001-004-01. We acknowledge the support of GPU cluster built by MCC Lab of Information Science and Technology Institution, USTC.

REFERENCES

- [1] Sk Miraj Ahmed, Dripta S Raychaudhuri, Sujoy Paul, Samet Oymak, and Amit K Roy-Chowdhury. 2021. Unsupervised multi-source domain adaptation without access to source data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 10103–10112.
- [2] Cosmin Ancuti, Codruța O Ancuti, Radu Timofte, and Christophe De Vleeschouwer. 2018. I-HAZE: a dehazing benchmark with real hazy and haze-free indoor images. In *International Conference on Advanced Concepts for Intelligent Vision Systems*. Springer, 620–631.
- [3] Dana Berman, Shai Avidan, et al. 2016. Non-local image dehazing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 1674–1682.
- [4] Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao. 2016. Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing* 25, 11 (2016), 5187–5198.
- [5] Chen Chen, Minh N Do, and Jue Wang. 2016. Robust image and video dehazing with visual artifact suppression via gradient residual minimization. In *Proceedings of the European Conference on Computer Vision (ECCV)*. Springer, 576–591.
- [6] Zeyuan Chen, Yangchao Wang, Yang Yang, and Dong Liu. 2021. PSD: Principled synthetic-to-real dehazing guided by physical priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 7180–7189.
- [7] Boris Chidlovskii, Stéphane Clinchant, and Gabriela Csurka. 2016. Domain adaptation in the absence of source domain data. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 451–460.
- [8] Qili Deng, Ziling Huang, Chung-Chi Tsai, and Chia-Wen Lin. 2020. Hardgan: A haze-aware representation distillation gan for single image dehazing. In *Proceedings of the European Conference on Computer Vision (ECCV)*. Springer, 722–738.
- [9] Hang Dong, Jinshan Pan, Lei Xiang, Zhe Hu, Xinyi Zhang, Fei Wang, and Ming-Hsuan Yang. 2020. Multi-scale boosted dehazing network with dense feature fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2157–2167.
- [10] Jiangxin Dong and Jinshan Pan. 2020. Physics-based feature dehazing networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*. Springer, 188–204.
- [11] Xinjie Fan, Qifei Wang, Junjie Ke, Feng Yang, Boqing Gong, and Mingyuan Zhou. 2021. Adversarially adaptive normalization for single domain generalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 8208–8217.
- [12] Raanan Fattal. 2008. Single image dehazing. *ACM Transactions on Graphics (TOG)* 27, 3 (2008), 1–9.
- [13] Raanan Fattal. 2014. Dehazing using color-lines. *ACM Transactions on Graphics (TOG)* 34, 1 (2014), 1–14.
- [14] Kaiming He, Jian Sun, and Xiaoou Tang. 2010. Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, 12 (2010), 2341–2353.
- [15] Tong He, Zhi Zhang, Hang Zhang, Zhongyue Zhang, Junyuan Xie, and Mu Li. 2019. Bag of tricks for image classification with convolutional neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 558–567.
- [16] Ming Hong, Yuan Xie, Cuihua Li, and Yanyun Qu. 2020. Distilling image dehazing with heterogeneous task imitation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 3462–3471.
- [17] Jiaxing Huang, Dayan Guan, Aoran Xiao, and Shijian Lu. 2021. Model adaptation: Historical contrastive learning for unsupervised domain adaptation without source data. *Advances in Neural Information Processing Systems* 34 (2021).
- [18] Xun Huang and Serge Belongie. 2017. Arbitrary style transfer in real-time with adaptive instance normalization. In *Proceedings of the IEEE International Conference on Computer Vision*. 1501–1510.
- [19] Songhao Jia, Ding-Jie Chen, and Hwann-Tzong Chen. 2019. Instance-level meta normalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 4865–4873.
- [20] Xin Jin, Cuiling Lan, Wenjun Zeng, Zhibo Chen, and Li Zhang. 2020. Style normalization and restitution for generalizable person re-identification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 3143–3152.
- [21] Boyun Li, Yuanbiao Gou, Shuhang Gu, Jerry Zitao Liu, Joey Tianyi Zhou, and Xi Peng. 2021. You only look yourself: Unsupervised and untrained single image dehazing neural network. *International Journal of Computer Vision* 129, 5 (2021), 1754–1767.
- [22] Boyun Li, Yuanbiao Gou, Jerry Zitao Liu, Hongyuan Zhu, Joey Tianyi Zhou, and Xi Peng. 2020. Zero-shot image dehazing. *IEEE Transactions on Image Processing* 29 (2020), 8457–8466.
- [23] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. 2017. Aod-net: All-in-one dehazing network. In *Proceedings of the IEEE International Conference on Computer Vision*. 4770–4778.
- [24] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. 2018. Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing* 28, 1 (2018), 492–505.
- [25] Lerenhan Li, Yunlong Dong, Wenqi Ren, Jinshan Pan, Changxin Gao, Nong Sang, and Ming-Hsuan Yang. 2019. Semi-supervised image dehazing. *IEEE Transactions on Image Processing* 29 (2019), 2766–2779.
- [26] Rui Li, Qianfen Jiao, Wenming Cao, Hau-San Wong, and Si Wu. 2020. Model adaptation: Unsupervised domain adaptation without source data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 9641–9650.
- [27] Runde Li, Jinshan Pan, Zechao Li, and Jinhui Tang. 2018. Single image dehazing via conditional generative adversarial network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 8202–8211.
- [28] Siyuan Li, Iago Breno Araujo, Wenqi Ren, Zhangyang Wang, Eric K Tokuda, Roberto Hirata Junior, Roberto Cesar-Junior, Jiawan Zhang, Xiaojie Guo, and Xiaochun Cao. 2019. Single image deraining: A comprehensive benchmark analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 3838–3847.
- [29] Risheng Liu, Xin Fan, Minjun Hou, Zhiying Jiang, Zhongxuan Luo, and Lei Zhang. 2018. Learning aggregated transmission propagation networks for haze removal and beyond. *IEEE Transactions on Neural Networks and Learning Systems* 30, 10 (2018), 2973–2986.
- [30] Xiaohong Liu, Yongrui Ma, Zhihao Shi, and Jun Chen. 2019. Griddehazenet: Attention-based multi-scale network for image dehazing. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 7314–7323.
- [31] Yang Liu, Jinshan Pan, Jimmy Ren, and Zhixun Su. 2019. Learning deep priors for image dehazing. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2492–2500.
- [32] Yuang Liu, Wei Zhang, and Jun Wang. 2021. Source-free domain adaptation for semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 1215–1224.
- [33] Ye Liu, Lei Zhu, Shunda Pei, Huazhu Fu, Jing Qin, Qing Zhang, Liang Wan, and Wei Feng. 2021. From Synthetic to Real: Image Dehazing Collaborating with Unlabeled Real Data. In *Proceedings of the 29th ACM International Conference on Multimedia*. 50–58.
- [34] Earl J McCartney. 1976. Optics of the atmosphere: scattering by molecules and particles. *New York* (1976).
- [35] Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik. 2012. No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing* 21, 12 (2012), 4695–4708.
- [36] Anish Mittal, Rajiv Soundararajan, and Alan C Bovik. 2012. Making a “completely blind” image quality analyzer. *IEEE Signal Processing Letters* 20, 3 (2012), 209–212.
- [37] Alan V Oppenheim and Jae S Lim. 1981. The importance of phase in signals. *Proc. IEEE* 69, 5 (1981), 529–541.
- [38] Xingang Pan, Ping Luo, Jianping Shi, and Xiaoou Tang. 2018. Two at once: Enhancing learning and generalization capacities via ibn-net. In *Proceedings of the European Conference on Computer Vision (ECCV)*. 464–479.
- [39] Zhihong Pan, Baopu Li, Dongliang He, Mingde Yao, Wenhao Wu, Tianwei Lin, Xin Li, and Errui Ding. 2022. Towards Bidirectional Arbitrary Image Rescaling: Joint Optimization and Cycle Idempotence. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 17389–17398.
- [40] Stephen M Pizer, E Philip Amburn, John D Austin, Robert Cromartie, Ari Geselowitz, Trey Greer, Bart ter Haar Romeny, John B Zimmerman, and Karel Zuiderweld. 1987. Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing* 39, 3 (1987), 355–368.
- [41] Xu Qin, Zhihui Wang, Yuanchao Bai, Xiaodong Xie, and Huizhu Jia. 2020. FFA-Net: Feature fusion attention network for single image dehazing. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 11908–11915.
- [42] Yanyun Qu, Yizi Chen, Jingying Huang, and Yuan Xie. 2019. Enhanced pix2pix dehazing network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 8160–8168.
- [43] Wenqi Ren, Si Liu, Hua Zhang, Jinshan Pan, Xiaochun Cao, and Ming-Hsuan Yang. 2016. Single image dehazing via multi-scale convolutional neural networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*. Springer, 154–169.
- [44] Wenqi Ren, Lin Ma, Jiawei Zhang, Jinshan Pan, Xiaochun Cao, Wei Liu, and Ming-Hsuan Yang. 2018. Gated fusion network for single image dehazing. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3253–3261.
- [45] Ali M Reza. 2004. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. *Journal of Signal Processing Systems for Signal Image and Video Technology* 38, 1 (2004), 35–44.
- [46] Yuanjie Shao, Lerenhan Li, Wenqi Ren, Changxin Gao, and Nong Sang. 2020. Domain adaptation for image dehazing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2808–2817.
- [47] Pranjay Shyam, Kuk-Jin Yoon, and Kyung-Soo Kim. 2021. Towards domain invariant single image dehazing. *arXiv preprint arXiv:2101.10449* (2021).
- [48] Nikolay Skarbnik, Yehoshua Y Zeevi, and Chen Sagiv. 2009. *The importance of phase in image processing*. Technion-Israel Institute of Technology, Faculty of Electrical Engineering.

- [49] Rosaura G VidalMata, Sreya Banerjee, Brandon RichardWebster, Michael Albright, Pedro Davalos, Scott McCloskey, Ben Miller, Asong Tambo, Sushobhan Ghosh, Sudarshan Nagesh, et al. 2020. Bridging the gap between computational photography and visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43, 12 (2020), 4272–4290.
- [50] Chao Wang, Hao-Zhen Shen, Fan Fan, Ming-Wen Shao, Chuan-Sheng Yang, Jian-Cheng Luo, and Liang-Jian Deng. 2021. EAA-Net: A novel edge assisted attention network for single image dehazing. *Knowledge-Based Systems* 228 (2021), 107279.
- [51] Menglu Wang, Xueyang Fu, Zepei Sun, and Zheng-Jun Zha. 2021. JPEG artifacts removal via compression quality ranker-guided networks. In *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*. 566–572.
- [52] Haiyan Wu, Yanyun Qu, Shaohui Lin, Jian Zhou, Ruizhi Qiao, Zhizhong Zhang, Yuan Xie, and Lizhuang Ma. 2021. Contrastive Learning for Compact Single Image Dehazing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 10551–10560.
- [53] Wenhan Yang, Ya Yuan, Wenqi Ren, Jiaying Liu, Walter J Scheirer, Zhangyang Wang, Taiheng Zhang, Qiaoyong Zhong, Di Xie, Shiliang Pu, et al. 2020. Advancing image understanding in poor visibility environments: A collective benchmark study. *IEEE Transactions on Image Processing* 29 (2020), 5737–5752.
- [54] Mingde Yao, Zhiwei Xiong, Lizhi Wang, Dong Liu, and Xuejin Chen. 2019. Spectral-depth imaging with deep learning based reconstruction. *Optics Express* 27, 26 (2019), 38312–38325.
- [55] Tian Ye, Mingchao Jiang, Yunchen Zhang, Liang Chen, Erkang Chen, Pen Chen, and Zhiyong Lu. 2021. Perceiving and Modeling Density is All You Need for Image Dehazing. *arXiv preprint arXiv:2111.09733* (2021).
- [56] Tian Ye, Yun Liu, Yunchen Zhang, Sixiang Chen, and Erkang Chen. 2022. Mutual Learning for Domain Adaptation: Self-distillation Image Dehazing Network with Sample-cycle. *arXiv preprint arXiv:2203.09430* (2022).
- [57] He Zhang and Vishal M Patel. 2018. Densely connected pyramid dehazing network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3194–3203.
- [58] Jing Zhang, Yang Cao, Shuai Fang, Yu Kang, and Chang Wen Chen. 2017. Fast haze removal for nighttime image using maximum reflectance prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 7418–7426.
- [59] Zhuoran Zheng, Wenqi Ren, Xiaochun Cao, Xiaobin Hu, Tao Wang, Fenglong Song, and Xiuyi Jia. 2021. Ultra-high-definition image dehazing via multi-guided bilateral learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 16180–16189.
- [60] Qingsong Zhu, Jiaming Mai, and Ling Shao. 2014. Single Image Dehazing Using Color Attenuation Prior.. In *British Machine Vision Conference(BMVC)*. Citeseer.
- [61] Qingsong Zhu, Jiaming Mai, and Ling Shao. 2015. A fast single image haze removal algorithm using color attenuation prior. *IEEE Transactions on Image Processing* 24, 11 (2015), 3522–3533.