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

Abstract

Deep learning-based methods have achieved remarkable performance for image dehazing. However, previous studies are mostly focused on training models with synthetic hazy images, which incurs performance drop when the models are used for real-world hazy images. We propose a Principled Synthetic-to-real Dehazing (PSD) framework to improve the generalization performance of dehazing. Starting from a dehazing model backbone that is pre-trained on synthetic data, PSD exploits real hazy images to fine-tune the model in an unsupervised fashion. For the fine-tuning, we leverage several well-grounded physical priors and combine them into a prior loss committee. PSD allows for most of the existing dehazing models as its backbone, and the combination of multiple physical priors boosts dehazing significantly. Through extensive experiments, we demonstrate that our PSD framework establishes the new state-of-the-art performance for real-world dehazing, in terms of visual quality assessed by no-reference quality metrics as well as subjective evaluation and downstream task performance indicator.

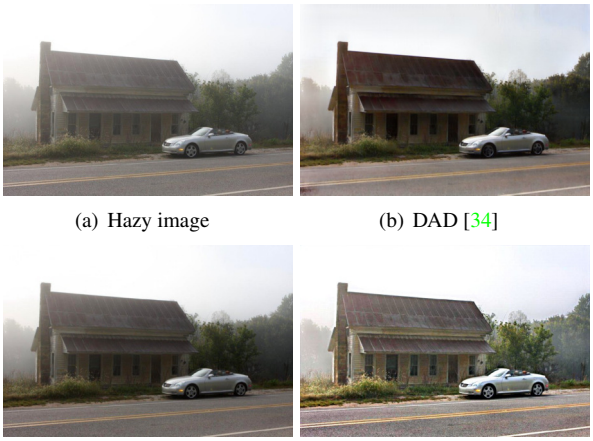
1. Introduction

Due to the existence of haze, outdoor images often suffer from low contrast and limited visibility, which adversely affects the performance of subsequent high-level computer vision tasks, such as object detection and recognition. Thus, more and more attention is drawn to image dehazing, that aims to recover the clean image from a hazy input.

According to the physical scattering model [24, 27], the hazing process is usually formulated as

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

*This work was supported by the Natural Science Foundation of China under Grants 62022075 and 61931014. Our code and models are available at https://github.com/zychen-ustc/PSD-Principled-Synthetic-to-Real-Dehazing-Guided-by-Physical-Priors. *(Corresponding author: Dong Liu.)*



(a) Hazy image (b) DAD [34] (c) MSBDN [6] (d) Ours

Figure 1. Dehazing results on a real hazy image.

where $I(x)$ is the observed hazy image, $J(x)$ is the scene radiance to be recovered. A and $t(x)$ are the global atmospheric light and the transmission map, respectively.

Estimating the clean image from a single hazy input is an ill-posed problem. Early methods [1, 8, 9, 11, 37] try to estimate transmission map by physical priors and then restore image via the scattering model. However, these physical priors are not always reliable, leading to inaccurate transmission estimates and unsatisfied dehazing results.

With the advances in deep learning, many methods based on convolutional neural networks (CNNs) [3, 4, 7, 14, 22, 31, 32] have been proposed to overcome the drawbacks of using physical priors. They are more efficient and outperform traditional prior-based algorithms. In common cases, large quantities of paired hazy/clean images are necessary for the training of CNN-based dehazing models. However, it is almost impossible to obtain these image pairs from the real world, and most learning-based methods resort to training on synthetic data. Unfortunately, due to the domain gap between synthetic and real data, dehazing models trained on synthetic images usually generalize poorly to real-world hazy images.

Recently, this issue has been picked up by a handful of

studies. Li *et al*. [17] proposed a semi-supervised dehazing model trained on both synthetic data and real-world images. Shao *et al*. [34] suggested a domain adaptation framework to reduce the gap between synthetic and real domains. These methods demonstrate the great potential of domain adaptation in improving the performance for real-world dehazing tasks. However, there is still room for further improvement, and principled frameworks dedicated for the problem of dehazing generalization remain a lack.

In this paper, we propose a Principled Synthetic-to-real Dehazing (PSD) framework, which is applicable to generalize most of the existing dehazing models to the real domain. PSD includes two phases: supervised pre-training and unsupervised fine-tuning. For pre-training, we modify a chosen dehazing model backbone into a physics-based network and train this network with synthetic data. Taking advantage of the well-designed backbone, we could obtain a pre-trained model with solid dehazing performance on the synthetic domain. For fine-tuning, we exploit real hazy images to train the model in an unsupervised manner. We investigate the strong physics background of dehazing tasks and elaborately select three physical priors to constitute a loss committee that guides the unsupervised training. The key idea of this fine-tuning process is the intuition that good dehazing results shall share some common properties/statistical priors. We leverage this intuition, treating the loss committee as a task-specific proxy guidance to help generalize our model to the real domain. As shown in Fig. 1, PSD produces a better dehazing result when compared with the backbone model. PSD also outperforms the state-of-art domain adaptation dehazing (DAD) [34].

We summarize the contributions of our work as follows.

- We re-formulate the real-world dehazing task as a *synthetic-to-real generalization* framework: starting from a dehazing model backbone pre-trained on synthetic paired data, real hazy images will be subsequently exploited to fine-tune the model in an unsupervised fashion. PSD is principled, easy-to-use, and can take most deep dehazing models as its backbone[1].

- Due to the absence of clean ground-truth image as supervision, we leverage several popular, well-grounded physical priors to guide the fine-tuning. We combine them into a *prior loss committee* as the task-specific proxy guidance, which constitutes the core of PSD. We show that these priors are complementary and their combination boosts PSD dehazing the most.

- Our framework is compared with a number of competitive methods via comprehensive experiments. Results are evaluated in terms of visual quality assessed by both no-reference quality metrics and subjective evaluation, and downstream task performance indicator. Consistently and substantially, PSD establishes the new state-of-the-art real-world dehazing performance.

[1]Some models are subject to light modifications, see Section 3.2.

2. Related Work

2.1. Single Image Dehazing

Prior-based methods. Dehazing methods based on priors [1, 8, 9, 11, 37, 43] estimate transmission maps by exploiting statistical properties of clean images, and then obtain dehazed results using the scattering model. In [37], Tan proposed a haze removal method by maximizing the local contrast of hazy images. He *et al*. [11] achieved impressive dehazing results using dark channel prior (DCP), which assumes that there exists at least one channel for every pixel whose value is close to zero. Zhu *et al*. [43] proposed a color attenuation prior to remove haze by estimating the scene depth. Fattal [9] introduced a color-line prior by the observation that pixels of small image patches typically exhibit a one-dimensional distribution in RGB color space. Berman *et al*. [1] proposed a dehazing algorithm based on the assumption that colors of a haze-free image are well approximated by a few hundred distinct colors. Although these methods have been shown effective for image dehazing, their performances are usually limited because these hand-crafted priors do not always hold for different hazy images.

Learning-based methods. With the availability of large-scale paired data and powerful CNNs, learning-based dehazing methods have become popular in recent years. MSCNN [31] is one of the first studies to solve haze removal problem via CNN, where the network is trained to estimate transmission map of the hazy input in a coarse-to-fine manner. Cai *et al*. [3] proposed an end-to-end network to generate transmission estimates. Zhang and Patel [42] embedded the physical scattering model into a network, which allows the network to estimate the transmission map, atmospheric light and dehazed image jointly. In addition, some other methods [6, 14, 18, 21, 28, 29, 32] have been proposed to directly recover the clean images. Li *et al*. [14] designed an AOD-Net to produce recovered images by reformulating the physical scattering model, and the same idea was later extended to video dehazing [15]. Ren *et al*. [32] introduced an gated fusion network which leverages the derived inputs from an original hazy image. Qu *et al*. [29] reformed the dehazing task into an image-to-image translation problem and proposed an enhanced pix2pix network to solve it. Dong *et al*. [6] incorporated a boosting strategy into the network to progressively restore clean image. All these methods have shown outstanding performances on dehazing.

However, the domain gap between synthetic and real data could cause a significant performance drop when these methods are generalized to real hazy images, since most of

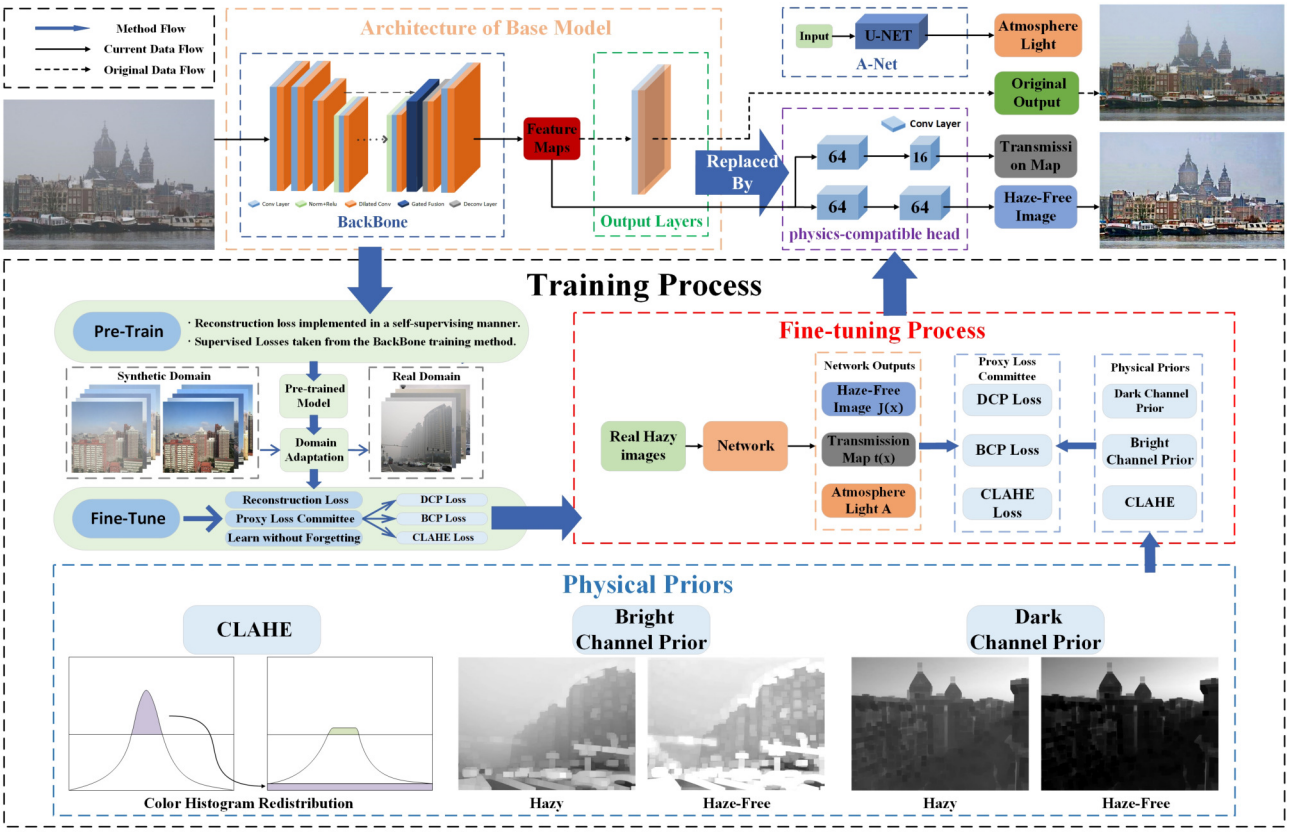


Figure 2. Overview of the proposed PSD framework. Our model consists of a backbone, a physics-compatible head, and an atmospheric light estimation network (A-net). We pre-train the model with synthetic images, then fine-tune the model with both synthetic and real hazy images, guided by a proxy loss committee derived from several physical priors.

them are only trained on synthetic data. To address this, we propose PSD, a principled synthetic-to-real dehazing generalization framework, intending to adapt synthetic data-based models to the real domain.

2.2. Unsupervised Domain Adaptation

Unsupervised domain adaptation aims to tackle domain shift between source and target domains, while images in the target domain are unlabeled. One major idea is to induce alignment between the source and target domains in feature space by optimizing for some measurements of distributional discrepancy [23, 35, 38]. Hoffman *et al*. [12] reduced the domain gap by using both generative image space alignment and latent representation space alignment. Zou *et al*. [44] proposed an unsupervised domain adaptation framework based on an iterative self-training procedure, where the loss of the latent variable is minimized.

Specifically, in dehazing and deraining tasks, there have been a few studies to solve the domain shift problem between synthetic and real domains [39,41]. Li *et al*. [17] proposed a semi-supervised learning (SSL) dehazing method, utilizing both synthetic and real images to learn features from both domains. Shao *et al*. [34] introduced a domain adaptation framework, which consists of an image translation module and two dehazing modules, one for the synthetic domain and the other for the real domain. Both methods show remarkable generalization performances. However, they do not make full use of the powerful prior knowledge in dehazing tasks and a performance gap between the two domains still exists. For this reason, we explore a large variety of physical priors and build up a loss committee as our proxy guidance for the training with real hazy images.

3. Proposed Method

3.1. Framework Overview

This section describes how PSD works. As shown in Fig. 2, PSD is a two-stage framework consisting of pre-training and fine-tuning.

Pre-training. We start by adopting one of the state-of-the-art dehazing models as our backbone, since these models achieve impressive performances on synthetic datasets and could implicitly provide the domain knowledge of hazy images. We then modify the backbone into a physics-based network that simultaneously generates the clean image $\tilde{J}$,

transmission map $\tilde{t}$, and atmospheric light $\tilde{A}$ from a single hazy input I. To jointly optimize the three components, we incorporate a reconstruction loss, which guides the network outputs to obey the physical scattering model (1).

In this phase, we only use the labeled synthetic data for training, and eventually obtain a model pre-trained on the synthetic domain.

Fine-tuning. For fine-tuning, we utilize unlabeled real data to generalize the pre-trained model from synthetic to real domain. Inspired by the strong physics background of dehazing, we believe that a high-quality haze-free image should obey some specific statistical rules, which can be derived from image priors. Furthermore, the physical knowledge provided by a single prior is not always reliable, so we aim to find a combination of multiple priors, hoping they can complement each other. Motivated by this, we design a prior loss committee as a task-specific proxy guidance for the training of unlabeled real data.

Besides, we implement a learning without forgetting (LwF) method [20], which forces our model to memorize the synthetic domain knowledge by forwarding the training data of the original task (*i.e.* synthetic hazy images) through the network together with real hazy data.

3.2. Physics-Based Network

Physical priors are usually related to the three components J, t, and A of the scattering model (1). As most deep dehazing models only estimate restored images directly, a modification on them is required. We propose to add two new modules.

Physics-compatible head. This module is composed of two branches, each of which contains two convolutional layers. We assume that the selected dehazing backbone is an effective feature extractor for both the transmission map t and the clean image J, and simple convolutional layers are enough to generate the two components from the feature maps. Therefore, we forward the backbone outputs through this physics-compatible head to produce a transmission map and a haze-free image separately, as shown in Fig. 2.

Atmospheric light estimation network (A-Net). A-Net is employed from DCPDN [42] and acts as an independent part of our model, which estimates the atmospheric light directly from a hazy input. We establish a connection between the backbone and this sub-network by a reconstruction loss, see Section 3.3 for details.

3.3. Model Pre-training

For the training of our modified network, we take the same parameter settings and the loss function $\mathcal{L}_o$ from the original backbone model. Since a single $\mathcal{L}_o$ loss is not able to update parameters of the two added modules, we include a reconstruction loss $\mathcal{L}_{Rec}$ into training to jointly optimize the entire model. Specifically, we aggregate the network outputs $\tilde{J}$, $\tilde{t}$, and $\tilde{A}$ to reconstruct the original input by the physical scattering model: $\tilde{I} = \tilde{J} \odot \tilde{t} + \tilde{A} \odot (1 - \tilde{t})$, where $\odot$ indicates element-wise multiplication. Then the reconstruction loss $\mathcal{L}_{Rec}$ is formulated as:

$$\mathcal{L}_{Rec} = \|I - \tilde{I}\|_1 \tag{2}$$

where I denotes the hazy input image.

The two losses $\mathcal{L}_o$ and $\mathcal{L}_{Rec}$ are combined for pre-training of the new physics-based network. Thanks to the well-designed backbone, our pre-trained model produces satisfactory dehazing results on synthetic hazy data.

3.4. Prior Loss Committee

We explore various dehazing priors, from which we select three effective and well-grounded ones. They provide us with the prior knowledge of real images. We combine the three priors into a loss committee as the task-specific proxy guidance for the unsupervised fine-tuning phase. The three members of this committee are described below in detail.

Dark Channel Prior (DCP) Loss. Dark Channel Prior (DCP) [11] is the most famous and effective prior for image dehazing. To implement DCP as a member of our prior loss committee, we follow the method proposed by [10] to reformulate this prior as an energy function:

$$\mathcal{L}_{DCP} = E(t, \tilde{t}) = t^T L t + \lambda (t - \tilde{t})^T (t - \tilde{t}) \tag{3}$$

where t and $\tilde{t}$ denote the transmission estimates from DCP and our network, respectively. L is a Laplacian-like matrix. The first term promotes successful image matting, and the second, fidelity to the dark channel solution. λ acts as a hyper-parameter.

Although $\mathcal{L}_{DCP}$ greatly advances the model performance on real hazy images, it generates side effects: the dehazing results are usually darker than expected. Thus, we integrate a bright channel prior loss into our committee as the second member.

Bright Channel Prior (BCP) Loss. Bright Channel Prior (BCP) is widely applied to solve dehazing and image enhancement problems. It helps make the resulting images brighter and with enhanced contrast. We implement BCP as the following loss function:

$$\mathcal{L}_{BCP} = \left\|t - \tilde{t}\right\|_1 \tag{4}$$

where t and $\tilde{t}$ represent the transmission estimates from BCP and our network, respectively.

$\mathcal{L}_{BCP}$ compensates the drawbacks caused by $\mathcal{L}_{DCP}$ through dramatically improving the global illumination of restored images and recovering more details. However, a committee with only two losses usually fails to maintain a stable fine-tuning process. In other words, it is not easy to achieve a balance between $\mathcal{L}_{DCP}$ and $\mathcal{L}_{BCP}$. Therefore, we enroll a new member into our prior loss committee, namely CLAHE reconstruction loss.

CLAHE Reconstruction Loss. Contrast Limited Adaptive Histogram Equalization (CLAHE) is a traditional contrast enhancement method, and it is also effective for image dehazing. Although we can restore hazy images directly using CLAHE, it is not advisable to treat its results as supervision, since it may bring the inherent flaws of this method into our network. Thus, we implement this prior in an indirect way. Specifically, we take two network outputs $\tilde{t}$, $\tilde{A}$, and the dehazing result of CLAHE J_{CLAHE} to reconstruct the original input via the scattering model (1) and then define the loss function by:

$$\mathcal{L}_{CLAHE} = \|I - I_{CLAHE}\|_1 \tag{5}$$

where I is the original hazy input, I_{CLAHE} is the reconstruction result by J_{CLAHE}, $\tilde{t}$, and $\tilde{A}$. $\mathcal{L}_{CLAHE}$ significantly improves the stability of our committee-guided unsupervised training process.

Finally, with all the three members, the prior loss committee provide a loss function defined as

$$\mathcal{L}_{com} = \lambda_d \mathcal{L}_{DCP} + \lambda_b \mathcal{L}_{BCP} + \lambda_c \mathcal{L}_{CLAHE} \tag{6}$$

where λ_d, λ_b, and λ_c are trade-off weights.

3.5. Synthetic-to-Real Generalization

With a pre-trained physics-based model $\mathcal{M}$ and a prior loss committee, we could step into the generalization phase, from synthetic to real. Starting from the model $\mathcal{M}$, we incorporate unlabeled real data into the training of $\mathcal{M}$ in an unsupervised fashion, by minimizing the loss function L_{com} provided by our prior loss committee. To avoid catastrophic forgetting, we also implement a learning without forgetting (LwF) loss $\mathcal{L}_{lwf}$, which helps our model memorize the previous dehazing task on synthetic datasets. Specifically, while updating the model $\mathcal{M}$ with real images, we keep a copy of the original model $\mathcal{M}_o$ which is frozen during the generalization process. We forward both synthetic and real images through $\mathcal{M}_o$ and minimize the difference between the output feature maps of $\mathcal{M}$ and $\mathcal{M}_o$. The loss function is formulated as:

$$\mathcal{L}_{lwf} = \|F_s - F_{os}\|_1 + \|F_r - F_{or}\|_1 \tag{7}$$

where F_s and F_{os} represent feature maps of $\mathcal{M}$ and $\mathcal{M}_o$ on synthetic data, F_r and F_{or} represent those on real data, respectively.

Besides, physical priors typically fail to correctly handle the sky in images, resulting in artifacts and color shifting. To address this issue, we roughly estimate the sky region of an input image by the dark channel prior and retain the original pixel values in the sky region as possible during fine-tuning, by the following loss function:

$$\mathcal{L}_{sky} = \|M_{sky} \odot (J - J_o)\|_1 \tag{8}$$

where M_{sky} is a binary mask indicating the sky region, J and J_o are the restored images from $\mathcal{M}$ and $\mathcal{M}_o$. More details are provided in the supplementary material.

The reconstruction loss $\mathcal{L}_{Rec}$ mentioned in Section 3.3 is also incorporated to integrate all the modules of the network and optimize them simultaneously.

Eventually, the overall loss function $\mathcal{L}$ in this phase is defined as

$$\mathcal{L} = \mathcal{L}_{com} + \mathcal{L}_{lwf} + \mathcal{L}_{sky} + \mathcal{L}_{Rec} \tag{9}$$

4. Experiments

4.1. Implementation Details

We choose OTS (Outdoor Training Set) and URHI (Unannotated Real Hazy Images) from the RESIDE dataset [16] for training, where synthetic images from OTS are for pre-training and real hazy images from URHI for fine-tuning. All the images are randomly cropped to patches of size 256×256, with normalized pixel values from -1 to 1.

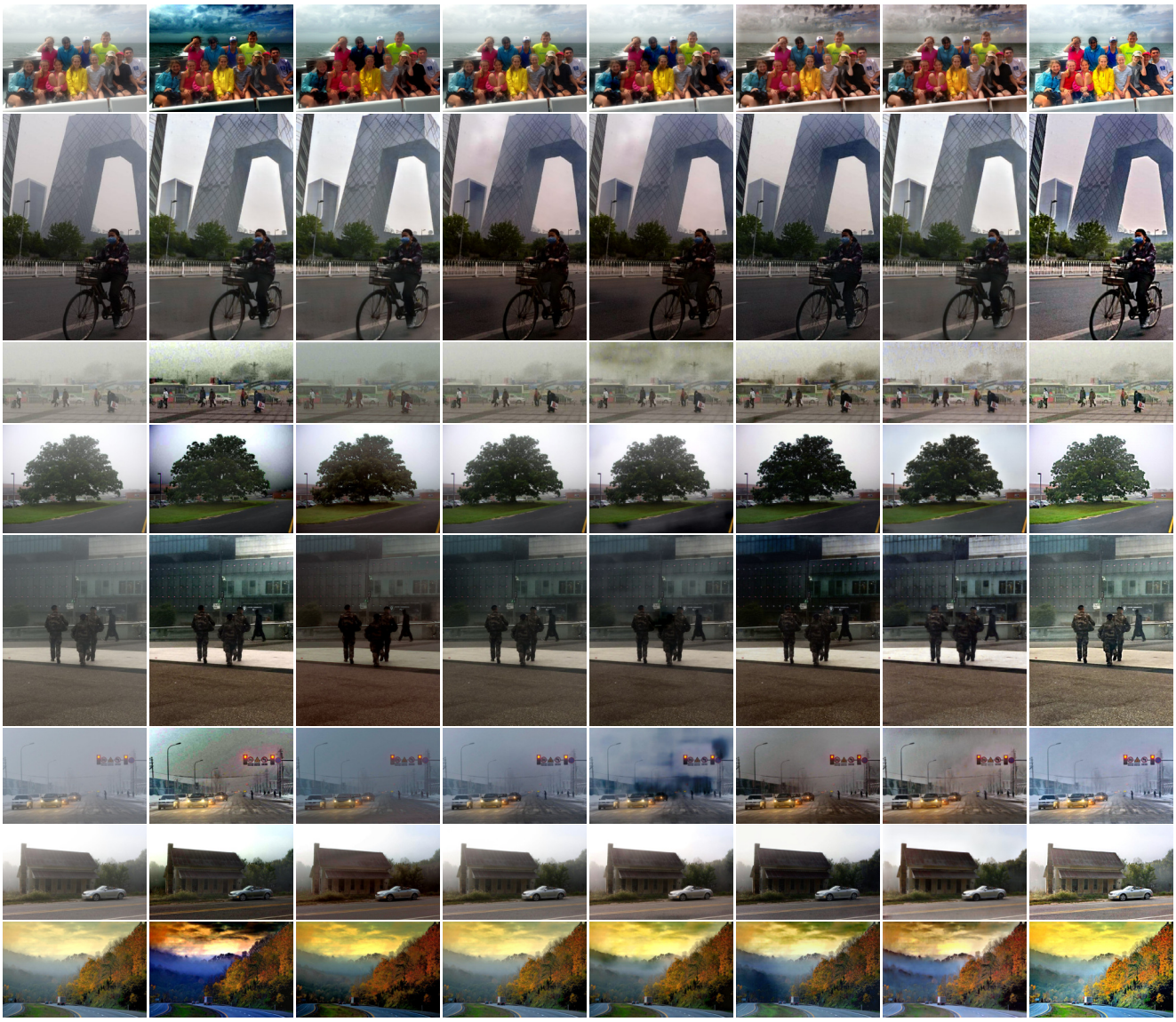
By default, the experiments of PSD are conducted on MSBDN [6] backbone, as this model provides state-of-the-art dehazing performance on synthetic images and is observed quite suitable for our framework. In pre-training, the backbone is modified and trained for 100 epochs by the Adam optimizer, with β_1 = 0.9 and β_2 = 0.999. The initial learning rate is set to 10^{-4} with a decay rate of 0.75 for every 10 epochs. In fine-tuning, we train the network for 20 epochs, with an initial learning rate set to 10^{-4} and decayed by 0.5 for every two epochs. The trade-off weights in loss function are set to $\lambda_d = 10^{-3}$, $\lambda_b = 0.05$, and $\lambda_c = 1$.

4.2. Comparison With State-of-the-Art Methods

We compare the performance of PSD with several state-of-the-art dehazing methods. A list of experiments are conducted, including visual quality comparison, human subjective survey, no-reference image quality assessment, and downstream task performance evaluation.

Visual Quality. We first evaluate the visual quality of PSD on real hazy images from RTTS, which is a subset of the RESIDE dataset [16]. We compare results of PSD with the following state-of-the-art methods: NLD [1], AOD-Net [14], FFANet [28], MSBDN [6], SSLD [17], EPDN [29], and DAD [34]. The results are shown in Fig. 3. *Except for images from RESIDE*, we also evaluate PSD on other real hazy images to further illustrate its superior generalization ability. These images are released by authors of previous studies [8, 9, 11]. The results are shown in Fig. 4.

From Figs. 3 and 4, we can observe that NLD [1] sometimes fails to handle the sky region, which leads to severe color shifting. Images restored by AOD-Net [14] and MSBDN [6] remain a bit hazy, especially in distant areas. SSLD [17] and EPDN [29] tend to darken the images, as



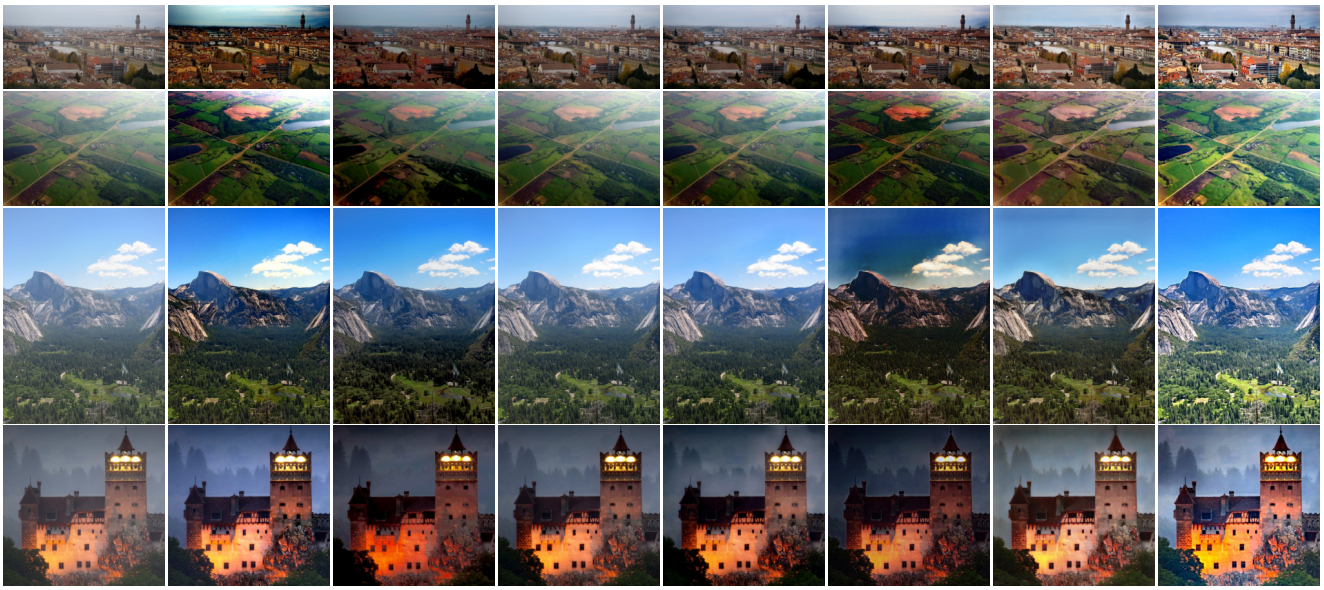
(a) Hazy images (b) NLD [1] (c) AOD-Net [14] (d) MSBDN [6] (e) SSLD [17] (f) EPDN [29] (g) DAD [34] (h) PSD

Figure 3. Comparison of dehazing results on real hazy images from RTTS.

shown in Fig. 3 (e) and Fig. 4 (f). The dehazing results of DAD [34] are pretty good, but still suffer from color distortion in some cases. Compared with all these methods, PSD generates high-quality haze-free images, with brighter details and sharper edges. More dehazing results are shown in the supplementary material.

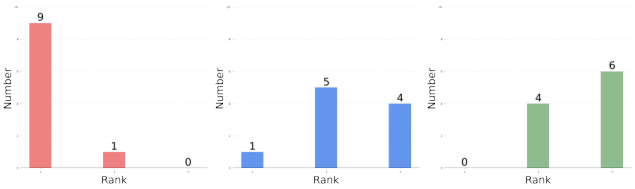
Human Subjective Evaluation. We conduct a human subjective study to evaluate the performance of PSD against other methods. We take 10 real-world hazy images from HSTS, a subset in RESIDE dedicated to subjective evaluation. We adopt PSD and the other two domain adaptation dehazing methods, SSLD [17] and DAD [34], to generate three dehazed images for each input. Following the protocol set by EnlightenGAN [13], we ask ten subjects to independently compare the results of the three methods in a pairwise manner. To be specific, each time a human subject is displayed with a pair of restored images randomly picked from the three results and is asked to choose an image that is more clean and visually pleasing, with high-quality details and no color shifting. As with EnlightenGAN, we fit a Bradley-Terry model [2] to estimate subjective scores of the methods. A rank in 1, 2, 3 is then assigned to the three methods on every image, according to the scores. Fig. 5 shows the evaluation results. PSD gets the highest score on 9 out of the 10 images, which indicates its superior generalization performance.

No-Reference Image Quality Assessment. For quantitative comparison, we use the Fog Aware Density Evalua-



(a) Hazy images (b) NLD [1] (c) AOD-Net [14] (d) MSBDN [6] (e) SSLD [17] (f) EPDN [29] (g) DAD [34] (h) PSD

Figure 4. Comparison of dehazing results on some real hazy images that were released by authors of previous methods.



(a) PSD (b) DAD [34] (c) SSLD [17]

Figure 5. Subjective evaluation results of three dehazing methods. In each histogram, x-axis denotes the rank (1–3, 1 represents the highest), and y-axis denotes the number of images of each rank. Our method scores the highest among 9 out of the 10 tested images.

Table 1. Quantitative results using NR-IQA metrics on RTTS. Red and blue indicate the best and the second-best, respectively.

Method	FADE↓	NIQE↓	BRISQUE↓	NIMA↑
Hazy	2.484	3.583	37.011	4.3250
MSBDN	1.363	3.154	28.743	4.1401
SSLD	0.867	3.489	32.428	4.2132
DAD	1.130	3.672	32.727	4.0055
PSD	0.920	3.077	25.239	4.3459

tor (FADE) [5] to assess the haze density of images. Also, we employ three well-known no-reference image quality assessment indicators: BRISQUE [25], NIQE [26], and NIMA [36]. All these metrics are evaluated on RTTS dataset. We compare PSD with SSLD, DAD, and MSBDN (the backbone). Results are reported in Table 1. Both BRISQUE and NIQE are blind image quality evaluators. They assess overall quality of images and PSD achieves the best. NIMA predicts aesthetic qualities of images, and PSD wins the first place again, which shows that the dehazing results of PSD are clean and visually pleasing. In FADE results, PSD comes the second. This is because FADE only focuses on the density of haze and ignores the degraded image details and color. The results of SSLD, as shown in Fig. 3 (e), are darker than others and hence look less hazy, which explains why it is preferred by FADE. However, SSLD removes many details and textures from the input images, while PSD achieves a better balance between dehazing cleanness and authenticity. In total, we win three of the four metrics, that further endorses the superiority of PSD on real-world dehazing tasks.

Task-Driven Evaluation. As pointed out by several recent studies [19,33,40], the performance of high-level computer vision tasks, such as object detection and recognition, will severely deteriorate in the presence of environmental degradations. Image dehazing could be used as a preprocessing step for these high-level tasks, and the resulting task performance could in turn become a downstream task performance indicator, as [13, 16, 19] suggested. We take RTTS as our test set, which consists of 4,322 real-world hazy images annotated with object categories and bounding boxes. We use YOLOv3 [30] for the object detection task and compute the mean Average Precision (mAP). All the dehazing models *only serve as pre-processing modules* to improve the quality of input images before the detection task, and no domain adaptation or joint learning processes are performed. The mAP just indicates the dehazing quality indirectly. Results are reported in Table 2. YOLO achieves the highest accuracy on the images processed by PSD, which provides a side evidence that PSD produces

(a) Hazy image (b) Without CLAHE loss (c) Without DCP loss (d) Without BCP loss (e) PSD

Figure 6. Results of the ablation study on the proxy loss committee.

(a) Hazy & DAD [34] (b) FFANet [28] (c) GridDehazeNet [21] (d) MSBDN [6]

Figure 7. Our PSD framework can choose different backbones. The first column shows the hazy image and its dehazing result using DAD [34]. Each column of (b, c, d) shows the dehazing result of one backbone and PSD upon that backbone.

Table 2. Object detection results on RTTS [16].

	Hazy	MSBDN	SSLD	DAD	PSD
mAP (%)	63.32	65.16	56.98	65.02	**65.84**
Gain	–	+1.84	-6.34	+1.70	**+2.52**

cleaner results and better preserves details.

4.3. Ablation Study

In order to demonstrate the effectiveness of our prior loss committee, we conduct an ablation study involving the following three control groups: 1) without DCP loss, 2) without BCP loss, 3) without CLAHE reconstruction loss. The visual results are shown in Fig. 6. In Fig. 6 (b), a committee without CLAHE reconstruction loss results in poor light uniformity of the restored image. In Fig. 6 (c), without DCP Loss, the image remains hazy, some image details are also blurred. Fig. 6 (d) shows that the absence of BCP Loss makes the dehazing result darker than expected. Finally, we can observe that a committee with all the three members generates the cleanest and most visually pleasing result, which indicates that the three priors complement each other.

4.4. Framework Generalization

We test two new backbones apart from the default MSBDN, namely FFANet [28] and GridDehazeNet [21], to show that PSD is generally applicable as a plug-and-play framework. We compare the backbones with their corresponding PSD models. In Fig. 7, PSD greatly improves the generalization of all the three backbones and achieves better results as compared to the latest domain adaptation dehazing (DAD) [34]. More images and the implementation details of new backbones are in the supplementary material.

5. Conclusion

We have presented a Principled Synthetic-to-real Dehazing (PSD) framework that is effective and easy to use. Starting from a backbone pre-trained on synthetic data, we use unlabeled real hazy images to fine-tune the network in an unsupervised fashion. This fine-tuning process is guided by a prior loss committee consisting of several popular and well-grounded physical priors. Extensive experiments demonstrate that PSD outperforms the state-of-the-art dehazing methods for real-world dehazing.

References

[1] Dana Berman, Tali Treibitz, and Shai Avidan. Non-local image dehazing. In *CVPR*, pages 1674–1682, 2016. 1, 2, 5, 6, 7

[2] Ralph Allan Bradley and Milton E. Terry. Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952. 6

[3] Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao. DehazeNet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 25(11):5187–5198, 2016. 1, 2

[4] Dongdong Chen, Mingming He, Qingnan Fan, Jing Liao, Liheng Zhang, Dongdong Hou, Lu Yuan, and Gang Hua. Gated context aggregation network for image dehazing and deraining. In *WACV*, pages 1375–1383. IEEE, 2019. 1

[5] Lark Kwon Choi, Jaehee You, and Alan Conrad Bovik. Referenceless prediction of perceptual fog density and perceptual image defogging. *IEEE Transactions on Image Processing*, 24(11):3888–3901, 2015. 7

[6] 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. 1, 2, 5, 6, 7, 8

[7] Jiangxin Dong and Jinshan Pan. Physics-based feature dehazing networks. In *ECCV*, pages 188–204. Springer, 2020. 1

[8] Raanan Fattal. Single image dehazing. *ACM Transactions on Graphics*, 27(3):1–9, 2008. 1, 2, 5

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

[10] Alona Golts, Daniel Freedman, and Michael Elad. Unsupervised single image dehazing using dark channel prior loss. *IEEE Transactions on Image Processing*, 29:2692–2701, 2019. 4

[11] 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. 1, 2, 4, 5

[12] Judy Hoffman, Eric Tzeng, Taesung Park, Jun-Yan Zhu, Phillip Isola, Kate Saenko, Alexei Efros, and Trevor Darrell. Cycada: Cycle-consistent adversarial domain adaptation. In *ICML*, pages 1989–1998. PMLR, 2018. 3

[13] Yifan Jiang, Xinyu Gong, Ding Liu, Yu Cheng, Chen Fang, Xiaohui Shen, Jianchao Yang, Pan Zhou, and Zhangyang Wang. EnlightenGAN: Deep light enhancement without paired supervision. *IEEE Transactions on Image Processing*, 30:2340–2349, 2021. 6, 7

[14] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. AOD-Net: All-in-one dehazing network. In *ICCV*, pages 4770–4778, 2017. 1, 2, 5, 6, 7

[15] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. End-to-end united video dehazing and detection. In *AAAI*, pages 7016–7023, 2018. 2

[16] 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. 5, 7, 8

[17] 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. 2, 3, 5, 6, 7

[18] Runde Li, Jinshan Pan, Zechao Li, and Jinhui Tang. Single image dehazing via conditional generative adversarial network. In *CVPR*, pages 8202–8211, 2018. 2

[19] 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. Single image deraining: A comprehensive benchmark analysis. In *CVPR*, pages 3838–3847, 2019. 7

[20] Zhizhong Li and Derek Hoiem. Learning without forgetting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(12):2935–2947, 2017. 4

[21] Xiaohong Liu, Yongrui Ma, Zhihao Shi, and Jun Chen. GridDehazeNet: Attention-based multi-scale network for image dehazing. In *ICCV*, pages 7314–7323, 2019. 2, 8

[22] Yang Liu, Jinshan Pan, Jimmy Ren, and Zhixun Su. Learning deep priors for image dehazing. In *ICCV*, pages 2492–2500, 2019. 1

[23] Mingsheng Long, Yue Cao, Jianmin Wang, and Michael Jordan. Learning transferable features with deep adaptation networks. In *ICML*, pages 97–105. PMLR, 2015. 3

[24] Earl J. McCartney. *Optics of the atmosphere–Scattering by molecules and particles*. Wiley, 1976. 1

[25] Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik. No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing*, 21(12):4695–4708, 2012. 7

[26] Anish Mittal, Rajiv Soundararajan, and Alan C. Bovik. Making a "completely blind" image quality analyzer. *IEEE Signal Processing Letters*, 20(3):209–212, 2012. 7

[27] Srinivasa G. Narasimhan and Shree K. Nayar. Vision and the atmosphere. *International Journal of Computer Vision*, 48(3):233–254, 2002. 1

[28] 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. 2, 5, 8

[29] Yanyun Qu, Yizi Chen, Jingying Huang, and Yuan Xie. Enhanced pix2pix dehazing network. In *CVPR*, pages 8160–8168, 2019. 2, 5, 6, 7

[30] Joseph Redmon and Ali Farhadi. YOLOv3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018. 7

[31] Wenqi Ren, Si Liu, Hua Zhang, Jinshan Pan, Xiaochun Cao, and Ming-Hsuan Yang. Single image dehazing via multi-scale convolutional neural networks. In *ECCV*, pages 154–169. Springer, 2016. 1, 2

[32] Wenqi Ren, Lin Ma, Jiawei Zhang, Jinshan Pan, Xiaochun Cao, Wei Liu, and Ming-Hsuan Yang. Gated fusion network for single image dehazing. In *CVPR*, pages 3253–3261, 2018. 1, 2

[33] Walter Scheirer, Rosaura VidalMata, Sreya Banerjee, et al. Bridging the gap between computational photography and visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, DOI:10.1109/TPAMI.2020.2996538*, 2020. 7

[34] Yuanjie Shao, Lerenhan Li, Wenqi Ren, Changxin Gao, and Nong Sang. Domain adaptation for image dehazing. In *CVPR*, pages 2808–2817, 2020. 1, 2, 3, 5, 6, 7, 8

[35] Baochen Sun and Kate Saenko. Deep CORAL: Correlation alignment for deep domain adaptation. In *ECCV*, pages 443–450. Springer, 2016. 3

[36] Hossein Talebi and Peyman Milanfar. NIMA: Neural image assessment. *IEEE Transactions on Image Processing*, 27(8):3998–4011, 2018. 7

[37] Robby T. Tan. Visibility in bad weather from a single image. In *CVPR*, pages 1–8. IEEE, 2008. 1, 2

[38] Eric Tzeng, Judy Hoffman, Ning Zhang, Kate Saenko, and Trevor Darrell. Deep domain confusion: Maximizing for domain invariance. *arXiv preprint arXiv:1412.3474*, 2014. 3

[39] Wei Wei, Deyu Meng, Qian Zhao, Zongben Xu, and Ying Wu. Semi-supervised transfer learning for image rain removal. In *CVPR*, pages 3877–3886, 2019. 3

[40] Wenhan Yang, Ye Yuan, Wenqi Ren, et al. Advancing image understanding in poor visibility environments: A collective benchmark study. *IEEE Transactions on Image Processing*, 29:5737–5752, 2020. 7

[41] Rajeev Yasarla, Vishwanath A. Sindagi, and Vishal M. Patel. Syn2Real transfer learning for image deraining using Gaussian processes. In *CVPR*, pages 2726–2736, 2020. 3

[42] He Zhang and Vishal M. Patel. Densely connected pyramid dehazing network. In *CVPR*, pages 3194–3203, 2018. 2, 4

[43] Qingsong Zhu, Jiaming Mai, and Ling Shao. A fast single image haze removal algorithm using color attenuation prior. *IEEE Transactions on Image Processing*, 24(11):3522–3533, 2015. 2

[44] Yang Zou, Zhiding Yu, Bvk Vijaya Kumar, and Jinsong Wang. Unsupervised domain adaptation for semantic segmentation via class-balanced self-training. In *ECCV*, pages 289–305, 2018. 3