

Cycle-Dehaze: Enhanced CycleGAN for Single Image Dehazing

Deniz Engin*

Anıl Genç*

Hazım Kemal Ekenel

SiMiT Lab, Istanbul Technical University, Turkey

{deniz.ingen, gencal6, ekenel}@itu.edu.tr

Abstract

In this paper, we present an end-to-end network, called Cycle-Dehaze, for single image dehazing problem, which does not require pairs of hazy and corresponding ground truth images for training. That is, we train the network by feeding clean and hazy images in an unpaired manner. Moreover, the proposed approach does not rely on estimation of the atmospheric scattering model parameters. Our method enhances CycleGAN formulation by combining cycle-consistency and perceptual losses in order to improve the quality of textural information recovery and generate visually better haze-free images. Typically, deep learning models for dehazing take low resolution images as input and produce low resolution outputs. However, in the NTIRE 2018 challenge on single image dehazing, high resolution images were provided. Therefore, we apply bicubic downscaling. After obtaining low-resolution outputs from the network, we utilize the Laplacian pyramid to upscale the output images to the original resolution. We conduct experiments on NYU-Depth, I-HAZE, and O-HAZE datasets. Extensive experiments demonstrate that the proposed approach improves CycleGAN method both quantitatively and qualitatively.

1. Introduction

Bad weather events such as fog, mist, and haze dramatically reduce the visibility of any scenery and constitute significant obstacles for computer vision applications, *e.g.* object detection, tracking, and segmentation. While images captured from hazy fields usually preserve most of their major context, they require some visibility enhancement as a pre-processing before feeding them into computer vision algorithms, which are mainly trained on the images captured at clear weather conditions. This pre-processing is generally called as *image dehazing/defogging*. Image dehazing techniques aim to generate haze-free images purified from the bad weather events. Sample hazy and haze-free



Figure 1: Hazy and clean examples from the NTIRE 2018 challenge on single image dehazing datasets: I-HAZE [6] & O-HAZE [7] datasets.

images from the NTIRE 2018 challenge on single image dehazing [4] are illustrated in Figure 1.

In the recent literature, researchers concentrate on single image dehazing methods, which can dehaze an input image without requiring any extra information, *e.g.* depth information or known 3D model of the scene. Single image dehazing approaches are divided into prior information-based methods [3, 5, 12, 16, 20, 32] and learning based methods [10, 24, 30, 34, 35]. Prior information-based methods are mainly based on the parameter estimation of atmospheric scattering model by utilizing the priors, such as dark channel priors [16], color attenuation prior [38], haze-line prior [8, 9]. On the other hand, these parameters are obtained from training data by learning based methods, which rely mostly on deep learning approaches. The proliferation of deep neural networks increases the use of large-scale datasets, therefore, researchers tend to create synthetic dehazing datasets like FRIDA [33] and D-HAZY [2], which have a more practical creation process than real dehazing datasets. Even though most of the deep learning approaches use the estimation of intermediate parameters, *e.g.* transmis-

*indicates equal contribution

sion map and atmospheric light [10, 24], there are also other approaches based on generative adversarial networks (GANs), which build a model without benefiting from these intermediate parameters [30].

GANs, introduced by Goodfellow *et al.* [14], are found to be very successful at image generation tasks, *e.g.* data augmentation, image inpainting, and style transfer. Their major goal is the generation of fake images indistinguishable from the original images on the targeted domain. By utilizing GANs, there exist *state-of-the-art* methods [30, 35] for single image dehazing, which require hazy input image and its ground truth in a paired manner. Recently, the need of paired data is removed after the cycle-consistency loss has been proposed by CycleGAN [37] for image-to-image translation. Inspired by the cycle-consistency loss, Disentangled Dehazing Network (DDN) has been introduced by Yang *et al.* for single image dehazing. Unlike CycleGAN [37] architecture, DDN reconstructs cyclic-image via the atmospheric scattering model instead of using another generator. Therefore, it requires the scene radiance, medium transmission map, and global atmospheric light [34] at the training phase.

In this work, we introduce Cycle-Dehaze network by utilizing CycleGAN [37] architecture via aggregating cycle-consistency and perceptual losses. Our main purpose is building an end-to-end network regardless of atmospheric scattering model for single image dehazing. In order to feed the input images into our network, they are resized to 256×256 pixel resolution via bicubic downscaling. After dehazing the input images, bicubic upscaling to their original size is not sufficient to estimate the missing information. To be able to obtain high-resolution images, we employed a simple upsampling method based on Laplacian pyramid. We perform our experiments on NYU-Depth [28] part of D-HAZY [2] and the NTIRE 2018 challenge on single image dehazing [4] datasets: I-HAZE [6] & O-HAZE [7]. According to our results, Cycle-Dehaze achieves higher image quality metrics than CycleGAN [37] architecture. Moreover, we analyze the performance of Cycle-Dehaze on cross-dataset scenarios, that is, we use different datasets at training and testing phases.

Our main contributions are summarized as follows:

- We enhance CycleGAN [37] architecture for single image dehazing via adding cyclic perceptual-consistency loss besides cycle-consistency loss.
- Our method requires neither paired samples of hazy and ground truth images nor any parameters of atmospheric scattering model during the training and testing phases.

- We present a simple and efficient technique to upscale dehazed images by benefiting from Laplacian pyramid.
- Due to its cyclic structure, our method provides a generalizable model demonstrated with the experiments on cross-dataset scenarios.

The rest of this paper is organized as follows: In Section 2, a brief overview of related work is provided. The proposed method is described in Section 3. Experimental results are presented and discussed in Section 4. Finally, the conclusions are given in Section 5.

2. Related Work

Image dehazing methods aim at recovering the clear scene reflection, atmospheric light color, and transmission map from an input hazy image. The requirement to know several parameters of the scene makes this problem challenging. Image dehazing methods can be categorized in terms of their inputs: (i) multiple images dehazing, (ii) polarizing filter-based dehazing, (iii) single image dehazing via utilizing additional information, *e.g.* depth or geometrical information methods, and (iv) single image dehazing [19].

Multiple images based methods overcome dehazing problem by obtaining changed atmospheric conditions from multiple images [21–23, 36]. In other words, it is required to wait until the weather condition or haze level are changed; thus, it is not practical for real-world applications. The polarization filter based approach has been proposed to dismiss the requirement of changed weather conditions [27]. In this approach, various filters are applied on different images to simulate changed weather conditions. Nevertheless, the static scenes are only considered when polarization filter based approaches are used. Therefore, this method still is not applicable for real-time scenarios. To address the necessities of these methods, single image dehazing via using additional information such as depth information [15] and the approximation of the 3D model of the scene [17] have been suggested. Since there is usually a single captured image of hazy scenes in the real-world conditions, obtaining additional information about the scene is extremely hard. Due to problems of previous approaches, researchers focus on single image dehazing methods.

Single Image Dehazing. Single image dehazing methods are mainly based on estimating parameters of the physical model, which is also known as the atmospheric scattering model. This model depends on the atmospheric condition of the scene, and can be expressed as follow:

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

where $I(x)$ is the hazy image, $J(x)$ is the haze-free image or the scene radiance, $t(x)$ is the medium transmission map, and A is the global atmospheric light on each x pixel coordinates. $t(x)$ can be formulated as:

$$t(x) = e^{-\beta d(x)} \quad (2)$$

where $d(x)$ is the depth of the scene point and β is defined as the scattering coefficient of the atmosphere.

Single image based methods can be categorized into two main approaches: prior information-based methods and learning-based methods. Prior information-based methods have been introduced as the pioneer of single image dehazing methods in [13, 31]. Following these studies, the dark channel prior (DCP) based on the statistics about the haze-free images has been presented by He *et al.* [16]. In this method, haze transmission map is estimated by utilizing dark pixels, which have a low-intensity value of at least one color channel. Dark channel prior has been enhanced by optimizing the inherent boundary constraint with weighted L1-norm contextual regularization to estimate transmission map [20]. In addition, Zhu *et al.* proposed a color attenuation prior (CAP) in order to recover depth information by creating a linear model on local priors [38]. Contrary to using local priors, Berman *et al.* introduced non-local color prior (NCP), which is based on an approximation of the *entire* haze-free images including few hundred distinct colors [8]. Each distinct color on a haze-free image is clustered and represented a line in RGB color space. Distance map and dehazed image are obtained by using these lines. Haze-line prior based approach has been improved by intersection with air-light to estimate global air-light in [9]. Moreover, due to non-uniform lighting conditions on the entire image, local airlight is estimated for each patch for night-time dehazing by multi-scale fusion in [3].

Recently, learning based methods have been employed by utilizing CNNs and GANs for single image dehazing. CNN based methods mainly focus on estimating transmission map and/or atmospheric light [8, 10] to recover clean images via atmospheric scattering model. On the other hand, GANs based methods produce haze-free images and estimate parameters of the physical model [34]. Also, combination of them has been proposed in [35].

Ren *et al.* [24] proposed a Multi-Scale CNN (MSCNN), which consists of coarse-scale and fine-scale networks in order to estimate the transmission map. The coarse-scale network estimates the transmission map, which is also improved locally by the fine-scale network. Another transmission map estimation network, called as DehazeNet, is designed differently from classical CNNs by adding feature extraction and non-linear regression layers, and it has been suggested by Cai *et al.* [10]. In addition to previous approaches, All-In-One Dehazing Network (AOD-net) has been presented in [18] to be able to produce clean images

directly without estimating intermediate parameters independently. Atmospheric scattering model is re-formulated to implement it in an end-to-end network. Zhang *et al.* [35] suggested a multi-task method that includes three modules, which are transmission map estimation via GANs, hazy feature extraction, and image dehazing. All modules have been trained jointly and image level loss functions, *e.g.* perceptual loss and pixel-wise Euclidean loss, have been utilized. Similarly, Yang *et al.* [34] introduced Disentangled Dehazing Network (DDN) to estimate the scene radiance, transmission map, and global atmosphere light by utilizing three generators jointly. Different from our method, these methods require estimation parameters of the atmospheric scattering model during training phase.

3. Proposed Method

Cycle-Dehaze is an enhanced version of CycleGAN [37] architecture for single image dehazing. In order to increase visual quality metrics, *PSNR*, *SSIM*, it utilizes the perceptual loss inspired by EnhanceNet [25]. The main idea of this loss is comparing images in a feature space rather than in a pixel space. Therefore, *Cycle-Dehaze* compares the original image with the reconstructed cyclic-image at both spaces, where cycle-consistency loss ensures a high *PSNR* value and perceptual loss preserves the sharpness of the image. Moreover, *Cycle-Dehaze* uses traditional Laplacian pyramid to provide better upsampling results after the main dehazing process. Figure 2 shows the overall representation of *Cycle-Dehaze* architecture.

As can be demonstrated in Figure 2, *Cycle-Dehaze* architecture is composed of two generators G, F and two discriminators D_x, D_y . In favor of cleaning/adding the haze, the architecture profits from the combination of cycle-consistency and cyclic perceptual-consistency losses besides the regular discriminator and generator losses. As a result of this, the architecture is forced to preserve textural information of the input images and generate unique haze-free outputs. On the other hand, pursuing the balance between cycle-consistency and perceptual-consistency losses is not a trivial task. Giving over-weight to perceptual loss causes the loss of color information after dehazing process. Therefore, cycle-consistency loss needs to have higher weights than the perceptual loss.

Cyclic perceptual-consistency loss. CycleGAN [37] architecture introduces cycle-consistency loss, which calculates $L1 - norm$ between the original and cyclic image for unpaired image-to-image translation task. However, this calculated loss between the original and cyclic image is not enough to recover all textural information, since hazy images are mostly heavily-corrupted. Cyclic perceptual-consistency loss aims to preserve original image structure by looking the combination of high and low-level features

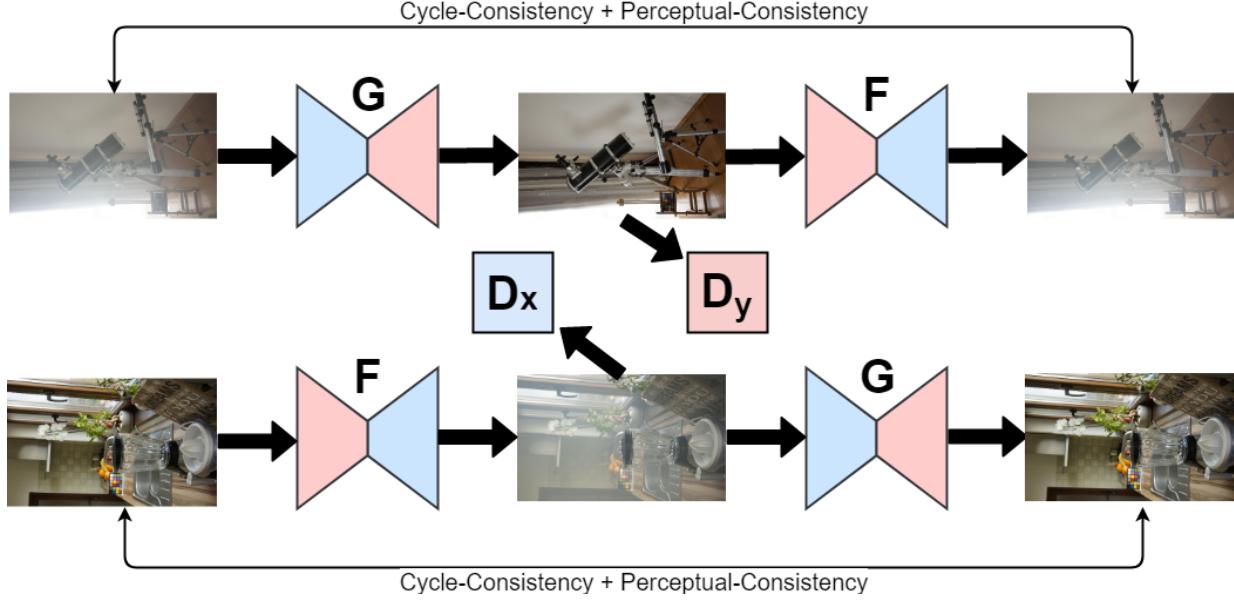


Figure 2: The architecture of Cycle-Dehaze Network where G & F refers to the generators, and D_x & D_y to the discriminators. For the sake of clarity, the representation is split into two parts: hazy to clean image, and clean to hazy image. **Best view in color.**

extracted from 2nd and 5th pooling layers of VGG16 [29] architecture. Under the constraints of $x \in X$, $y \in Y$ and generator $G : X \rightarrow Y$, generator $F : Y \rightarrow X$, the formulation of cyclic perceptual-consistency loss is given below, where (x, y) refers to hazy and ground truth unpaired image set and ϕ is a VGG16 [29] feature extractor from 2nd and 5th pooling layers:

$$\begin{aligned} \mathcal{L}_{Perceptual} = & \|\phi(x) - \phi(F(G(x)))\|_2^2 \\ & + \|\phi(y) - \phi(G(F(y)))\|_2^2. \end{aligned} \quad (3)$$

Full objective of Cycle-Dehaze. Cycle-Dehaze has one extra loss compared to CycleGAN [37] architecture. Therefore, the objective of Cycle-Dehaze can be formulated as follows, where $\mathcal{L}_{CycleGAN}(G, F, D_x, D_y)$ is the full objective of CycleGAN [37] architecture, D stands for the discriminator and γ controls the effect of cyclic perceptual-consistency loss:

$$\begin{aligned} \mathcal{L}(G, F, D_x, D_y) = & \mathcal{L}_{CycleGAN}(G, F, D_x, D_y) \\ & + \gamma * \mathcal{L}_{Perceptual}(G, F), \end{aligned} \quad (4)$$

$$G^*, F^* = \arg \min_{G, F, D_x, D_y} \max \mathcal{L}(G, F, D_x, D_y). \quad (5)$$

Conclusively, Cycle-Dehaze optimizes CycleGAN [37] architecture with the additional cyclic perceptual-consistency loss given in Equation 3 according to Equations 4 and 5. In order to obtain haze-free images, the generator G^* is used at the testing time.

Laplacian upscaling. Cycle-Dehaze architecture takes 256×256 pixel resolution input image and produces 256×256 pixel resolution output image because of GPU limitation. In order to reduce deterioration of the image quality during the downscaling and upscaling process, we have taken advantage of Laplacian pyramid, which is created by using high-resolution hazy images. In order to get the high-resolution dehazed image, we have changed the top layer of Laplacian pyramid with our dehazed low-resolution image and performed Laplacian upscaling process as usual. This basic usage of Laplacian pyramid especially preserves most of the edges of the hazy image during the cleaning process and boosts SSIM value at the upscaling stage. Laplacian upscaling is an optional post-processing step while working on the high-resolution images.

Implementation details (indoor/outdoor). We used TensorFlow [1] framework for the training and testing phases, and MATLAB for resizing images. We trained our model with NVIDIA TITAN X graphics card. We performed around 40 epochs on each dataset in order to ensure convergence. Our testing time is about 8 seconds per image on Intel Core i7-5820K CPU. During the training phase of our model, we used Adam optimizer with the learning rate $1e-4$. Moreover, we took γ as 0.0001 which is $1e-5$ times lower than the weight of the cycle-consistency loss.

Metrics	None	He <i>et al.</i> [16]	Zhu <i>et al.</i> [38]	Berman <i>et al.</i> [8]	Ren <i>et al.</i> [24]	Cai <i>et al.</i> [10]	CycleGAN	Yang <i>et al.</i> [34]	Ours
PSNR	9.46	10.98	12.78	12.26	13.04	12.84	13.38	15.54	15.41
SSIM	0.58	0.64	0.70	0.70	0.66	0.71	0.52	0.77	0.66

Table 1: Average PSNR and SSIM results on NYU-Depth [28] dataset. Most of the accuracies taken from the paper [34]. Numbers in red and blue indicate first and second best results, respectively. The second column of the table shows the values which are average PSNR and SSIM results calculated directly between the each hazy and its ground truth image.

Our network is similar to the original CycleGAN [37] architecture except for the cyclic perceptual-consistency loss and Laplacian pyramid as a post-processing step. At the highest level of Laplacian pyramid, we scale the images to 256×256 pixel resolution because of our network’s requirements. To calculate cyclic perceptual-consistency loss, we used VGG16 [29] architecture, which is initialized by ImageNet [11] pre-trained model. The source code of the proposed method will be publicly available through project’s GitHub page¹.

4. Experiments and Results

In this section, we have presented the experimental results and discussed them with the results of the challenge. The first experiment is to compare our result with the *state-of-the-art* approaches on NYU-Depth [28] dataset. Then, we have investigated our performance on the NTIRE 2018 challenge on single image dehazing [4] datasets: I-HAZE [6] & O-HAZE [7]. In addition, we have emphasized differences between CycleGAN [37] and our proposed method, *Cycle-Dehaze*, via qualitative and quantitative results. Furthermore, we have provided comparative qualitative results on natural images.

4.1. Datasets

NYU-Depth [28] dataset consists of 1449 pairs of clean and synthesized hazy images of the same scene. The dataset is the part of D-HAZY [2] dataset, which includes two individual environments presented as Middlebury [26] and NYU-Depth [28]. We have chosen NYU-Depth [28], which is considerably larger scale than Middlebury [26] part. NYU-Depth [28] contains also depth map of each scene, which is not used for this study.

The *NTIRE 2018 challenge on single image dehazing* [4] datasets were collected via professional fog generators and camera setup for image dehazing problem. Each image includes a Macbeth ColorChecker mostly used for color calibration with the real-world. The challenge [4] has two main datasets: I-HAZE [6] & O-HAZE [7], which have 25 indoor and 35 outdoor hazy images and their ground truth images, respectively. The captured images are in very high resolution. During the challenge, the organizers provide

additional 10 images for each dataset as a validation and test set. We did not include them in the training data.

Data Augmentation. We have employed data augmentation by taking random crops as the pre-processing step before the training phase. This procedure makes our model robust for different scales and textures. Our data augmentation procedure is as in Algorithm 1.

Algorithm 1 Data Augmentation

```

1: procedure AUGMENTER
2:   factor  $\leftarrow$  the number of augmented image
3:   Image crops[1..factor]  $\leftarrow$  []
4:   i  $\leftarrow$  0
5:   loop:
6:     if i = factor then
7:       WRITE(crops);
8:       break
9:     x, y  $\leftarrow$  select a random pixel coordinate on image
10:    w, h  $\leftarrow$  select a random width, height
11:    crops(i)  $\leftarrow$  CROP(x,y,w,h)
12:    crops(i)  $\leftarrow$  RESIZE(crops(i), [256, 256])
13:    i  $\leftarrow$  i + 1.
14:   goto loop.

```

According to Algorithm 1, we take random crops from the image by selecting random pixel coordinates and crop sizes. Then, we resize our crops to 256×256 before feeding them into our network. We run our *augmenter* function for each image in our datasets.

During data augmentation, we have obtained 200 images per original image in the training set of the I-HAZE [6] and O-HAZE [7] datasets, since this dataset contains too few images.

4.2. Results on NYU-Depth Dataset

We have conducted our experiments on the benchmark NYU-Depth [28] dataset to illustrate the performance of our approach compared to the other *state-of-the-art* approaches. We have employed Cycle-Dehaze by taking hazy images as input, and compared haze-free outputs with theirs ground truths. We have tested our method on all images of NYU-Depth [28] dataset and reported average PSNR and SSIM values in Table 1.

¹github.com/engindeniz/Cycle-Dehaze

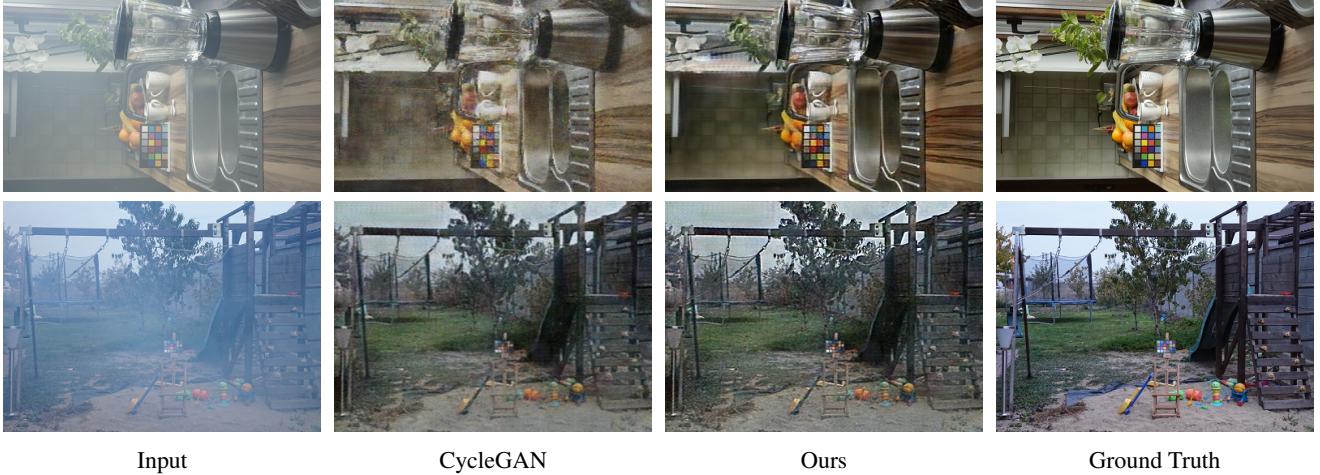


Figure 3: Qualitative results on I-HAZE [6] & O-HAZE [7] datasets.

Methods	Metrics	I-HAZE [6]		O-HAZE [7]	
		Validation	Test	Validation	Test
None	PSNR / SSIM	13.80 / 0.73	- / -	13.56 / 0.59	- / -
Best	PSNR / SSIM	20.41 / 0.85	24.97 / 0.88	22.82 / 0.74	24.59 / 0.77
CycleGAN	PSNR / SSIM	17.80 / 0.75	- / -	18.92 / 0.53	- / -
Ours	PSNR / SSIM	18.03 / 0.80	18.30 / 0.80	19.92 / 0.64	19.62 / 0.67

Table 2: Average PSNR and SSIM results on the NTIRE 2018 challenge on single image dehazing [4] datasets: I-HAZE [6] & O-HAZE [7]. According to preliminary results, the first row demonstrates the best accuracies of the NTIRE 2018 challenge on single image dehazing [4]. The first row of the results shows the values which are average PSNR and SSIM results calculated directly between the each hazy and its ground truth image.

Table 1 compares our quantitative results with the other approaches including CycleGAN [37]. By outperforming the *state-of-the-art* methods [8, 10, 16, 24, 38] according to PSNR metric, our model achieves the second best result. Moreover, Cycle-Dehaze reaches higher PSNR and SSIM values than CycleGAN [37]. This demonstrates that adding perceptual-consistency loss on CycleGAN [37] improves the architecture further for PSNR and SSIM metrics on NYU-Depth [28] dataset. The results also indicate that Cycle-Dehaze could get nearly similar PSNR results with the approaches profited from parameters of the atmospheric scattering model [34].

4.3. Results on I-HAZE and O-HAZE Datasets

We have focused on the NTIRE 2018 challenge on single image dehazing [4] datasets: I-HAZE [6] & O-HAZE [7] during the preparation of this work. We have analyzed effects of Laplacian pyramid and cyclic perceptual-loss, especially on I-HAZE [6] dataset. The challenge datasets are considerably higher resolution than other image dehazing datasets *e.g.* NYU-Depth [28]. Therefore, the scaling process on images has a larger effect on I-HAZE [6]

and O-HAZE [7] datasets according to PSNR and SSIM metrics. Our Laplacian pyramid reduces this deforming effect of the scaling process. We have tested our method on all validation and test images of the challenge datasets provided by organizers of the NTIRE 2018 challenge on single image dehazing [4]. We have trained our final model only on the training set, which is also provided by the organizers. Table 2 presents average PSNR and SSIM values and Figure 3 shows sample qualitative results.

Quantitative Results. According to Table 2, our proposed method gave better PSNR and SSIM values than CycleGAN [37] for each track of the challenge. This shows that additional cyclic perceptual-consistency loss and Laplacian pyramid increase the performance of original CycleGAN [37] architecture. Moreover, PSNR and SSIM differences between the I-HAZE [6] and O-HAZE [7] datasets presents that outdoor scenes suffer from SSIM values because of the long shot of the captured images. On the other hand, they have higher PSNR values since the produced fog spreads the atmosphere and the captured images seem less hazed than the indoor scenes.

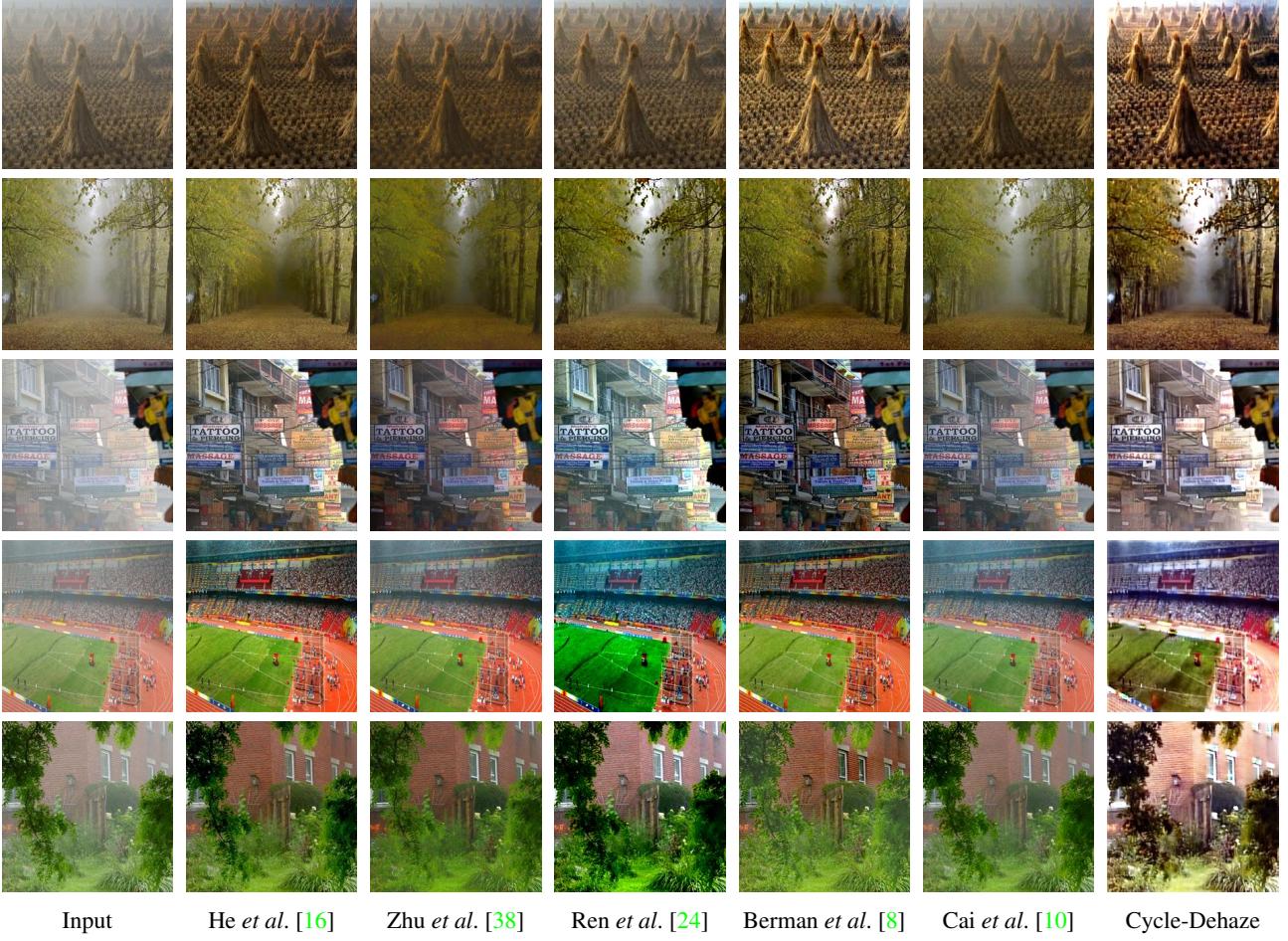


Figure 4: Qualitative results on natural hazy images by comparing with state-of-the-art-results.

Qualitative Results. Figure 3 shows the qualitative difference between CycleGAN [37] and Cycle-Dehaze. From the qualitative results, it can be clearly seen that dehazed images by Cycle-Dehaze has less noise and sharper edges for both indoor and outdoor scenes, where cyclic perceptual-consistency loss reduces the noise of the dehazed images and Laplacian pyramid leads sharper edges. Since outdoor scenes include more textural repetitions than indoor scenes, recovering textures of O-HAZE [7] is harder than I-HAZE [6]. Therefore, our sharpness on edges reduces in outdoor conditions.

4.4. Results on Cross-Dataset Image Dehazing

CNNs mostly tend to overfit on a specific dataset rather than learning the targeted task. To the best of our knowledge, fine-tuning the trained model on a targeted dataset is the most popular solution of overfitting. On the other hand, we have analyzed our method with two distinct experiments in cross-dataset setups, in which entirely different datasets have been used for the training and testing phases. Firstly,

we have tested Cycle-Dehaze on some popular natural images used by image dehazing community by scaling them to 256×256 . Figure 4 provides the qualitative results obtained on them. Secondly, we have tested the final model trained on NYU-Depth [28] dataset on I-HAZE [6] dataset and vice versa, since both of the datasets have been created under indoor conditions. Table 3 presents the accuracies of cross-dataset testing and Figure 5 shows the visual difference of cross-dataset testing between the datasets captured at indoor scenes: NYU-Depth [28] and I-HAZE [6] datasets.

Training set	Test set	PSNR	SSIM
NYU-Depth [28]	NYU-Depth [28]	15.41	0.66
I-HAZE [6]	NYU-Depth [28]	13.12	0.59
I-HAZE [6]	I-HAZE [6]	18.03	0.80
NYU-Depth [28]	I-HAZE [6]	14.76	0.73

Table 3: Cross-dataset quantitative results of Cycle-Dehaze architecture on the datasets captured at indoor scenes.

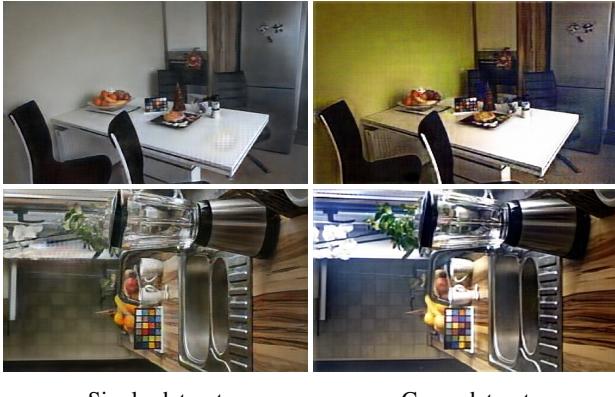


Figure 5: Comparative qualitative results between single and cross dataset experiments via Cycle-Dehaze.

Quantitative Results. According to Table 3, Cycle-Dehaze obtains considerably high PSNR and SSIM values on cross-dataset testing. As a matter of fact, the results both on NYU-Depth [28] and I-HAZE [6] are as fine as original CycleGAN [37] architecture on regular single dataset testing. This shows that Cycle-Dehaze mostly learns the dehazing task rather than overfitting on a dataset. Due to the cyclic mechanism of Cycle-Dehaze, our method focus on adding a haze on images beside cleaning a haze. Therefore, Cycle-Dehaze learns what is haze regardless of the image dehazing problem. On the other side, the methods only addressed the dehazing process tend to focus on color enhancement on the specific dataset. From the result of cross-dataset experiments, Cycle-Dehaze can be considered as a practical solution on real-world conditions for single image dehazing.

Qualitative Results. Figure 4 shows the comparative qualitative results of Cycle-Dehaze on natural hazy images with respect to state-of-the-art image dehazing methods [8, 10, 16, 24, 38]. According to qualitative results, the performance of Cycle-Dehaze is perceptually satisfying on natural images, especially when the color tones of neighboring pixels are very close to each other. Specifically, Cycle-Dehaze preserves the natural color toning of hazy image after dehazing process. Consequently, Cycle-Dehaze keeps the shadows and depth on the image more perceptible.

Figure 5 includes the images, which are dehazed by regular Cycle-Dehaze and by the cross-dataset version of it. According to qualitative results, both methods can clear the haze from the input images. On the other hand, the color recovery on the single dataset is better than on cross-dataset which leads lower PSNR results on cross-dataset scenario. Since the haze thickens at some parts of the images, our model can not estimate the actual ground truth color if it is trained on another dataset.

5. Conclusion

We proposed a single image dehazing network, named as Cycle-Dehaze, which directly generates haze-free images from hazy input images without estimating parameters of the atmospheric scattering model. Besides, our network provides a training process of hazy and ground truth images in an unpaired manner. In order to retain the high visual quality of haze-free images, we improved cycle-consistency loss of CycleGAN architecture by combining it with the perceptual loss. Cycle-Dehaze takes low-resolution images as input, so it requires downscaling of its inputs as a pre-processing step. For reducing distortion on images while resizing, we utilized Laplacian pyramid to upscale low-resolution images instead of using directly bicubic upscaling. The experimental results show that our method produces visually better images and achieves higher PSNR and SSIM values than CycleGAN architecture. Moreover, we performed additional experiments on the cross-dataset scenario to demonstrate generalizability of our model for different domains.

Acknowledgements. We would like to thank our colleagues from SiMiT Lab at ITU and LTS5 at EPFL, especially Christophe René Joseph Ecabert and Saleh Bagher Salimi, for their valuable comments. Travel grant for this research is provided by Yapı Kredi Technology.

References

- [1] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.
- [2] C. Ancuti, C. O. Ancuti, and C. De Vleeschouwer. D-HAZY: a dataset to evaluate quantitatively dehazing algorithms. In *IEEE International Conference on Image Processing (ICIP)*, pages 2226–2230. IEEE, 2016.
- [3] C. Ancuti, C. O. Ancuti, C. De Vleeschouwer, and A. C. Bovik. Night-time dehazing by fusion. In *IEEE International Conference on Image Processing (ICIP)*, pages 2256–2260. IEEE, 2016.
- [4] C. Ancuti, C. O. Ancuti, R. Timofte, L. Van Gool, L. Zhang, M.-H. Yang, et al. NTIRE 2018 challenge on image dehazing: Methods and results. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2018.
- [5] C. O. Ancuti, C. Ancuti, C. Hermans, and P. Bekaert. A fast semi-inverse approach to detect and remove the haze from a single image. In *Asian Conference on Computer Vision*, pages 501–514. Springer, 2010.
- [6] C. O. Ancuti, C. Ancuti, R. Timofte, and C. De Vleeschouwer. I-HAZE: a dehazing benchmark with real hazy and haze-free indoor images. In *arXiv*, 2018.

- [7] C. O. Ancuti, C. Ancuti, R. Timofte, and C. De Vleeschouwer. O-HAZE: a dehazing benchmark with real hazy and haze-free outdoor images. In *arXiv*, 2018.
- [8] D. Berman, S. Avidan, et al. Non-local image dehazing. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1674–1682, 2016.
- [9] D. Berman, T. Treibitz, and S. Avidan. Air-light estimation using haze-lines. In *IEEE International Conference on Computational Photography (ICCP)*, pages 1–9. IEEE, 2017.
- [10] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao. Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 25(11):5187–5198, 2016.
- [11] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.
- [12] S. Emberton, L. Chittka, and A. Cavallaro. Hierarchical rank-based veiling light estimation for underwater dehazing.
- [13] R. Fattal. Single image dehazing. *ACM Transactions on Graphics (TOG)*, 27(3):72, 2008.
- [14] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems (NIPS)*, pages 2672–2680, 2014.
- [15] N. Hautière, J.-P. Tarel, and D. Aubert. Towards fog-free in-vehicle vision systems through contrast restoration. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8. IEEE, 2007.
- [16] K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12):2341–2353, 2011.
- [17] J. Kopf, B. Neubert, B. Chen, M. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, and D. Lischinski. *Deep photo: Model-based photograph enhancement and viewing*, volume 27. ACM, 2008.
- [18] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng. Aod-net: All-in-one dehazing network. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 4770–4778, 2017.
- [19] Y. Li, S. You, M. S. Brown, and R. T. Tan. Haze visibility enhancement: A survey and quantitative benchmarking. *Computer Vision and Image Understanding*, 165:1–16, 2017.
- [20] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan. Efficient image dehazing with boundary constraint and contextual regularization. In *IEEE International Conference on Computer Vision (ICCV)*, pages 617–624. IEEE, 2013.
- [21] S. G. Narasimhan and S. K. Nayar. Chromatic framework for vision in bad weather. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 598–605. IEEE, 2000.
- [22] S. G. Narasimhan and S. K. Nayar. Contrast restoration of weather degraded images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(6):713–724, 2003.
- [23] S. K. Nayar and S. G. Narasimhan. Vision in bad weather. In *IEEE International Conference on Computer Vision (ICCV)*, volume 2, pages 820–827. IEEE, 1999.
- [24] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang. Single image dehazing via multi-scale convolutional neural networks. In *European Conference on Computer Vision*, pages 154–169. Springer, 2016.
- [25] M. S. Sajjadi, B. Schölkopf, and M. Hirsch. Enhancenet: Single image super-resolution through automated texture synthesis. In *IEEE International Conference on Computer Vision (ICCV)*, 2017.
- [26] D. Scharstein, H. Hirschmüller, Y. Kitajima, G. Krathwohl, N. Nešić, X. Wang, and P. Westling. High-resolution stereo datasets with subpixel-accurate ground truth. In *German Conference on Pattern Recognition*, pages 31–42. Springer, 2014.
- [27] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar. Instant dehazing of images using polarization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages I–I. IEEE, 2001.
- [28] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus. Indoor segmentation and support inference from rgbd images. In *European Conference on Computer Vision*, pages 746–760. Springer, 2012.
- [29] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations (ICLR)*, 2015.
- [30] K. Swami and S. K. Das. Candy: Conditional adversarial networks based fully end-to-end system for single image haze removal. *arXiv preprint arXiv:1801.02892*, 2018.
- [31] R. T. Tan. Visibility in bad weather from a single image. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8. IEEE, 2008.
- [32] J.-P. Tarel and N. Hautiere. Fast visibility restoration from a single color or gray level image. In *IEEE International Conference on Computer Vision*, pages 2201–2208. IEEE, 2009.
- [33] J.-P. Tarel, N. Hautiere, L. Caraffa, A. Cord, H. Halmaoui, and D. Gruyer. Vision enhancement in homogeneous and heterogeneous fog. *IEEE Intelligent Transportation Systems Magazine*, 4(2):6–20, 2012.
- [34] X. Yang, Z. Xu, and J. Luo. Towards perceptual image dehazing by physics-based disentanglement and adversarial training. In *The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)*, 2018.
- [35] H. Zhang, V. Sindagi, and V. M. Patel. Joint transmission map estimation and dehazing using deep networks. *arXiv preprint arXiv:1708.00581*, 2017.
- [36] T. Zhang, C. Shao, and X. Wang. Atmospheric scattering-based multiple images fog removal. In *4th International Congress on Image and Signal Processing (CISP)*, volume 1, pages 108–112. IEEE, 2011.
- [37] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. *IEEE International Conference on Computer Vision (ICCV)*, 2017.
- [38] Q. Zhu, J. Mai, and L. Shao. A fast single image haze removal algorithm using color attenuation prior. *IEEE Transactions on Image Processing*, 24(11):3522–3533, 2015.