

Enhanced CycleGAN Dehazing Network

Zahra Anvari and Vassilis Athitsos

VLM Research Lab, Computer Science and Engineering Department,
University of Texas at Arlington, Texas, U.S.A.

Keywords: Image Restoration and Reconstruction, Single Image Dehazing, Generative Adversarial Network.

Abstract: Single image dehazing is a challenging problem, and it is far from solved. Most current solutions require paired image datasets that include both hazy images and their corresponding haze-free ground-truth. However, in reality lighting conditions and other factors can produce a range of haze-free images that can serve as ground truth for a hazy image, and a single ground truth image cannot capture that range. This limits the scalability and practicality of paired methods in real-world applications.

In this paper, we focus on *unpaired* single image dehazing and reduce the image dehazing problem to an unpaired image-to-image translation and propose an Enhanced CycleGAN Dehazing Network (ECDN). We enhance CycleGAN from different angles for the dehazing purpose. We employ a global-local discriminator structure to deal with spatially varying haze. We define self-regularized color loss and utilize it along with perceptual loss to generate more realistic and visually pleasing images. We use an encoder-decoder architecture with residual blocks in the generator with skip connections so that the network better preserves the details. Through an ablation study, we demonstrate the effectiveness of different modules in the performance of the proposed network. Our extensive experiments over two benchmark datasets show that our network outperforms previous work in terms of PSNR and SSIM.

1 INTRODUCTION

Haze is an atmospheric phenomenon that can cause visibility issues, and the quality of images captured under haze can be severely degraded. Hazy images suffer from poor visibility and low contrast, which can challenge both human visual perception and numerous intelligent systems relying on computer vision methods.

The performance of standard computer vision tasks such as object detection (Liu et al., 2016; Redmon et al., 2016), semantic segmentation (Long et al., 2015), face detection, clustering and dataset creation (Yang et al., 2016; Anvari and Athitsos, 2019; Lin et al., 2018; Lin et al., 2017) can be affected significantly when images are hazy. Hence, image dehazing is an essential pre-processing task for general-purpose computer vision algorithms that are fed with hazy images. As a result, single image dehazing has received a great deal of attention over the past decade (Ancuti et al., 2016; Ancuti et al., 2010; Emberton et al., 2015; Meng et al., 2013; Tarel and Hautiere, 2009; Ancuti et al., 2016; Ancuti et al., 2010; Emberton et al., 2015; Meng et al., 2013; Anvari and Athitsos, 2020; Tarel and Hautiere, 2009).

Most of the recent image dehazing methods rely on paired datasets, which means for each hazy image there's a single clean/haze-free image as a ground truth. In practice, however, there is a range of clean images that can correspond to a hazy image, due to factors such as contrast or light intensity changes throughout the day. In fact, it is infeasible to capture both ground truth/clear image and the hazy image of the same scene simultaneously. Thus there is an emerging need to develop solutions that do not rely on the ground truth images and could operate with *unpaired* supervision.

Single image dehazing methods can be categorized into two main classes: prior-based methods and learning-based methods. Prior-based models solve the haze removal problem through estimating the physical model, *i.e.* transmission map and atmospheric light parameters. Learning-based methods mainly use CNN-based or GAN-based models to recover the haze-free images. These models take advantage of large amount of training data to learn a model that recovers the haze-free image of a hazy image.

In this paper, we focus on *unpaired* image dehazing and first cast the unpaired image dehazing problem to an image-to-image translation problem and

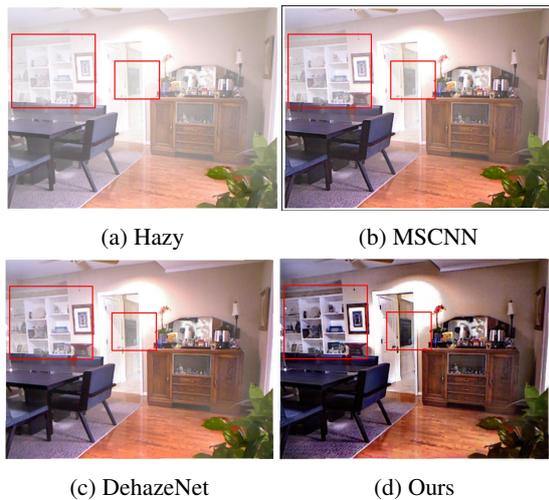


Figure 1: A single image dehazing example. Our method generates an image with less haze and rich details compared with MSCNN and DehazeNet.

then propose a novel cycle-consistent generative adversarial network, called ECDN, that operates without paired supervision and benefits from (i) a global-local discriminator architecture to handle spatially varying haze (ii) an encoder-decoder generator architecture with residual blocks to better preserve the details (iii) skip connections in the generator to improve the performance of the network and convergence (iv) customized cyclic perceptual loss and a self-regularized color loss to generate more realistic images and mitigate the color distortion problem. Through empirical analysis we show that the proposed network can effectively remove haze and generate visually pleasing haze-free images.

Figure 1 shows the result of our method compared to the current state-of-the-art methods. Our proposed method removes haze more effectively and generates a more realistic clean image compared to previous work.

In summary, this paper presents the following contributions:

- We propose a novel cycle-consistent generative adversarial network called ECDN for unpaired image dehazing. ECDN does not rely on any priors such as the physical scattering model, as opposed to many previous methods, and instead it adopts the image-to-image translation approach for unpaired image dehazing.
- We adopt a global-local discriminator structure to deal with spatially varying haze and generate better haze-free images.
- We define a self-regularized color loss and utilize it along with a customized perceptual loss to gen-

erate more visually pleasing images with vibrant colors and mitigate the color distortion problem. Self-regularization is vital to our network since in unpaired setting there is no external supervision available.

- We use an encoder-decoder generator architecture with residual blocks with skip connections to better preserve the details.
- Through empirical analysis, we show that our network outperforms the previous work in terms of PSNR and SSIM.

2 RELATED WORK

Numerous attempts have been done to solve the single image haze removal problem. These methods can be categorized into two main classes: prior-based and learning-based, that we describe them below.

2.1 Prior-based Dehazing

Prior-based methods are mainly based on prior information and assumptions to recover the haze-free images from hazy images. They heavily depend on estimating the parameters of the physical scattering model (McCartney, 1976; Srinivasa and Shree, 2002), *aka.* the atmospheric scattering model, which contains the transmission map and the atmospheric light to solve the haze removal problem. The physical scattering model is formulated as:

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

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. He *et al.* (He et al., 2010) proposed a dark channel prior to estimate the transmission map effectively. Tan et al. (Tan, 2008) increase the contrast of hazy images, based on the fact that haze-free images have higher contrast than hazy images.

2.2 Learning-based Dehazing

Recently learning based methods have been proposed that utilize CNNs and GANs for the single image dehazing problem. CNN-based methods try to recover the clean images through the atmospheric scattering model, by mainly estimating the transmission map and atmospheric light (McCartney, 1976; Narasimhan and Nayar, 2000).

MSCNN (Ren et al., 2016) contains two sub-networks called coarse-scale and fine-scale, to estimate the transmission map. The coarse-scale network estimates the transmission map and is further improved locally by the fine-scale network. In DehazeNet (Cai et al., 2016), authors modified the classic CNN model by adding feature extraction and non-linear regression layers. These modifications distinguish DehazeNet from other CNN-based models. The All-In-One Dehazing Network (AOD-net) (Li et al., 2017) proposed an end-to-end network that produces the haze-free/clean images through reformulating the atmospheric scattering model.

2.3 Generative Adversarial Networks

GANs have become one of the most successful methods for image generation, manipulation, restoration, and reconstruction. GANs have been used to super-resolve images (Ledig et al., 2017), remove motion blurriness from images (Kupyn et al., 2018), and remove noise (Chen et al., 2018), to name a few applications. GANs are also utilized in image dehazing. DDN was proposed as a disentangled dehazing network without paired supervision (Yang et al., 2018). The GAN that they proposed consists of three generators: one for generating haze-free image, one for the atmospheric light, and the third generator for transmission map.

The Cycle-consistent GAN (CycleGAN) (Zhu et al., 2017) method was proposed for unpaired image-to-image translation task and has gained significant attention during the past couple of years. CycleGAN is utilized for image dehazing along with the perceptual loss to generate more visually realistic dehazed images (Engin et al., 2018).

3 PROPOSED METHOD

First we reduce the unpaired image dehazing problem to an image-to-image translation problem, and then propose an Enhanced CycleGAN Dehazing Network (ECDN) to translate a hazy image to a haze-free one. Next we describe our network in details.

3.1 Overview of ECDN

Figure 2a demonstrates an overview of our proposed network. On the left, you can see the two domains *i.e.* hazy and haze-free, and the generator G_A which generates haze-free image of a hazy image and G_B which does the backward translation from haze-free to hazy.

We need these forward and backward translations to ensure the cycle consistency. At each direction we have two discriminators *i.e.* D_{global} and D_{local} for each generator to enforce them to generate more realistic and better haze-free images.

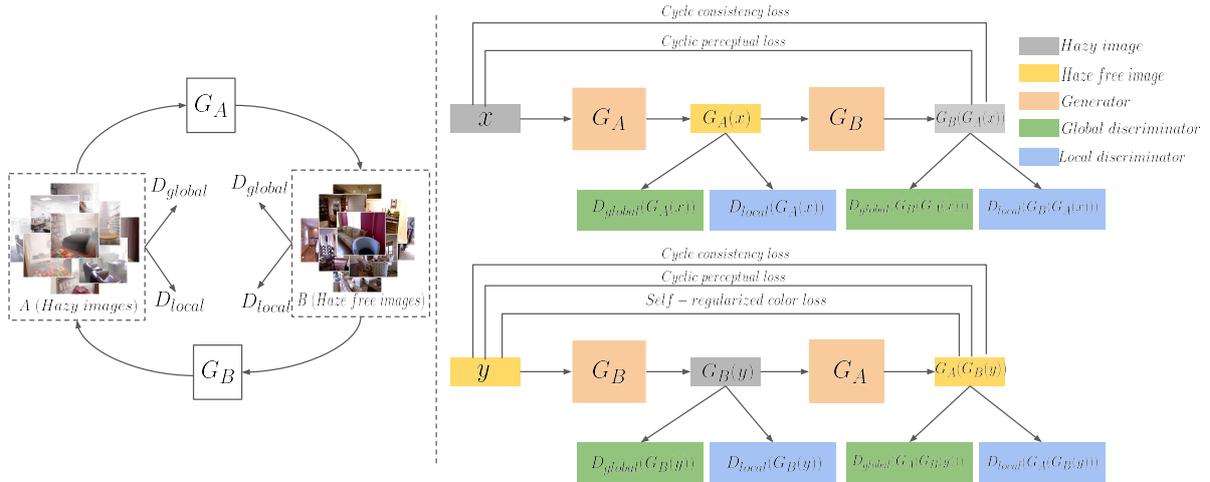
Right side of Figure 2a illustrates our proposed network in forward and backward cycles. Top row depicts the hazy to haze-free translation cycle and how the components interact. x is the hazy image and $G_B(G_A(x))$ is the reconstructed hazy image that is used to calculate loss values *i.e.* cycle consistency loss and cyclic perceptual loss. The bottom row shows the backward cycle *i.e.* how the haze-free image is reconstructed through the backward cycle. y is the haze-free image and $G_A(G_B(y))$ is the reconstructed haze-free image that is used to calculate different loss values *i.e.* cycle consistency loss, cyclic perceptual loss, and also self-regularized color loss. We only use self-regularized color loss in the backward cycle, since we want to make the haze free and the reconstructed haze free images closer in terms of color, and prevent color shifting and distortion.

Figure 2b depicts the network architecture of the generator G_A and the global and local discriminators. G_A and G_B utilize the same network architecture. Similarly all discriminators share the same network architecture, however operate on different scales.

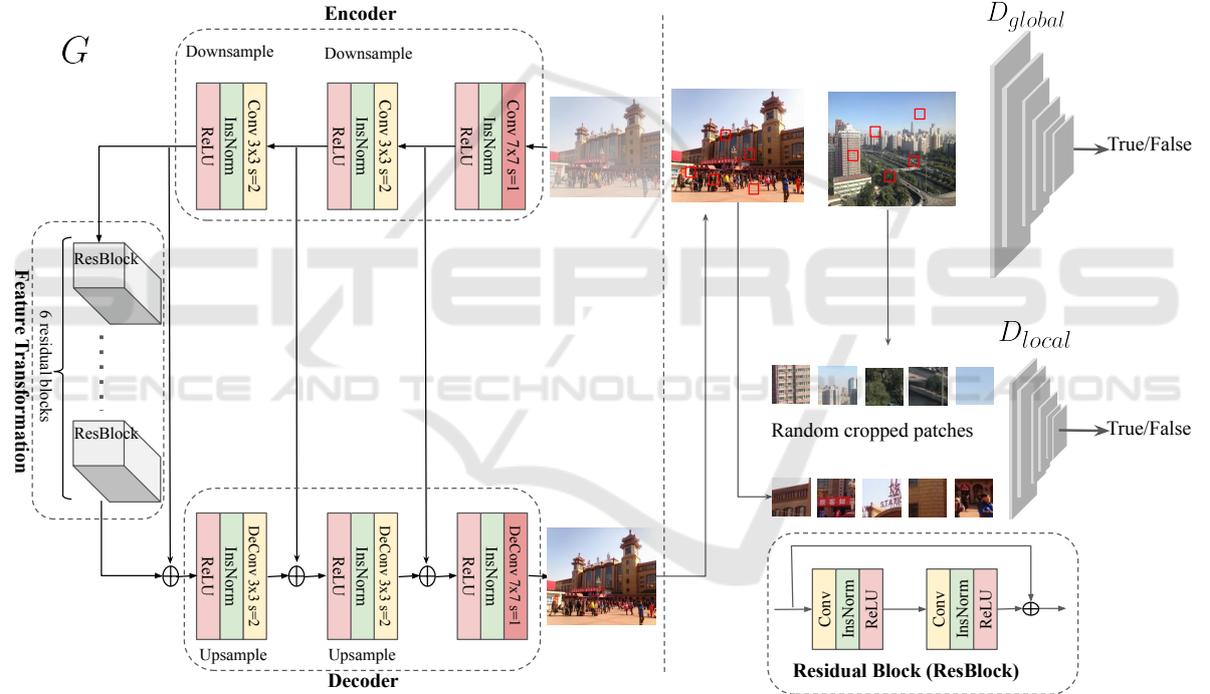
3.2 Generator

Figure 2b presents the architecture of ECDN model. The architecture of generator G_A is depicted on the left. Note that G_B has the same architecture as G_A . In order to generate a haze-free image without paired supervision in a cycle-consistent manner, we require a generator network that can preserve the images' texture, structure and details while removing haze. Therefore, we designed a network with three parts: encoder, feature transformation, and decoder.

The encoder module starts with a convolution layer followed by an Instance Normalization and Relu non-linearity and two downsampling blocks. Feature transformation, has six Residual Blocks to extract complex and deep features whilst removing haze. Going deeper in network helps it to become capable of representing complex functions and also learn features at many different levels of abstraction. Decoder consists of two upsampling blocks which are deconvolution layers, followed by Instance Normalization and Relu. The deconvolution layers are used to recover image structural details and convert the feature maps to a haze-free RGB image. The upsampling operations are performed through the deconvolution layer to obtain intermediate feature mappings with



(a) An overview of ECDN



(b) The architecture of ECDN. This figure shows the architecture of G_A , D_B^{Global} and D_B^{Local} . G_B , D_A^{Global} and D_A^{Local} have the same architecture as G_A , D_B^{Global} , D_B^{Local} respectively, except that they work on different inputs, *i.e.*, the input to G_B is a clean image and the input to G_A is a hazy image.

Figure 2: The overview and architecture of ECDN.

double spatial size and half channels than its previous counterpart.

We use skip links between corresponding layers of different levels from encoder and decoder to guarantee better convergence. A skip connection before downsampling, is also applied between input and output of the feature transformation module, as shown in Figure 2b.

3.3 Discriminator

The right side of Figure 2b shows D_B^{Global} and D_B^{Local} . Note that D_A^{Global} and D_A^{Local} have the same architecture as D_B^{Global} and D_B^{Local} respectively. We have two types of discriminators, global and local, each performing a particular operation to classify real vs. fake images. Initially our model contained only global dis-

criminator. However, we have observed that global discriminators often fail on spatially-varying hazy images, *i.e.*, in cases where haze density variation exists in an image. Thus we decided that different image parts need to be enhanced differently. In order to enhance each region of an image appropriately, in addition to improving the haze removal globally, we utilized a global-local discriminator scheme inspired by (Jiang et al., 2019) in a cycle-consistent manner.

Global discriminator D_B^{Global} classifies if a haze-free image generated by G_A is real or fake, **based on the entire image**. Local discriminator D_B^{Local} classifies if a haze-free image generated by G_A is real or fake, **based on 5 randomly cropped image patches of size 64×64 pixels from that image**.

3.4 Loss Functions

Our objective loss function contains:

- **Adversarial Loss** for matching the distribution of generated images to the data distribution in the target domain.
- **Cycle Consistency Loss** to prevent the learned mappings G_A and G_B from contradicting each other.
- **Cyclic Perceptual Loss** to help the generators generate more realistic and visually pleasing images.
- **Self-regularized Color Loss** to avoid color shifting and artifacts in generated haze-free images and also guide the generator to generate images with vibrant colors.

The overall loss function for training ECDN is defined as follows:

$$Loss_{total} = L_{global}^{GAN} + L_{local}^{GAN} + L_{global}^{Cycle} + L_{local}^{Cycle} + L_{global}^{CP} + L_{local}^{CP} + L_{global}^{SRColor} \quad (2)$$

Next we describe these loss functions in details.

3.4.1 Adversarial Loss

We adopted Least Squares GAN to calculate the adversarial loss. Equations 3 and 4 show how we calculate the adversarial loss for the global discriminators and the global generators respectively.

$$L_D^{Global} = E_{x_r \sim P_{real}} [(D(x_r) - 1)^2] + E_{x_f \sim P_{fake}} [(D(x_f) - 0)^2] \quad (3)$$

$$L_G^{Global} = E_{x_r \sim P_{fake}} [(D(x_r) - 1)^2] \quad (4)$$

where D denotes the discriminator, and x_r and x_f are sampled from the real and fake distribution respectively.

We introduced the local discriminator to further enhance hazy image and deal with spatially-varying hazy images. Equations 5 and 6 depicts the corresponding loss functions:

$$L_D^{Local} = E_{x_r \sim P_{real-patches}} [(D(x_r) - 1)^2] + E_{x_f \sim P_{fake-patches}} [(D(x_f) - 0)^2] \quad (5)$$

$$L_G^{Local} = E_{x_f \sim P_{fake-patches}} [(D(x_f) - 1)^2] \quad (6)$$

where D denotes the discriminator, x_r and x_f are sampled from **patches** taken from real and fake distributions.

3.4.2 Cycle Consistency Loss

Adversarial loss can not guarantee that the learned function can map an individual input x_i to desired output y_i . Thus a cycle-consistency loss is proposed by CycleGAN to reduce the space of possible mapping functions. Cycle-consistency loss function (*L1-norm*) compares the cyclic image and the original image in an unpaired image-to-image translation process (Zhu et al., 2017). Cycle consistency loss is defined as:

$$L_{cycle}(G_A, G_B) = E_{x \sim P_{data(x)}} [\|G_B(G_A(x)) - x\|_1] + E_{y \sim P_{data(y)}} [\|G_A(G_B(y)) - y\|_1] \quad (7)$$

where G_A and G_B are forward and backward generators, x belongs to domain X (*i.e.* the original domain, hazy images here) and y belongs to domain Y (*i.e.* the haze-free images). $G_B(G_A(x))$ and $G_A(G_B(y))$ are the reconstructed images.

3.4.3 Self-regularized Color Loss

Hazy images usually lack brightness and contrast, to improve these lacking features we define a self-regularized color loss, inspired by (Wang et al., 2019) to measure color difference between the haze-free images and the reconstructed images. We call it *self-regularized* because we do not rely on the ground truth image.

This loss function forces the generator to generate images with the same color distribution as the haze-free images. In addition, we observed that some of the reconstructed images have color artifacts which is an inherent problem of CycleGAN, this loss function was employed to deal with this problem as well. Equation 8 shows color loss function.

$$L_{SRColor} = \sum_p ANGLE(G_A(G_B(y))_p, y_p) \quad (8)$$

Where $()_p$ denotes a pixel; ANGLE is a function that calculates the angle between two colors regarding the RGB color as a 3D vector. y belongs to domain Y (*i.e.* haze-free images) and $G_A(G_B(y))$ the reconstructed haze-free image.

Eq. 8 sums the angles between the color vectors for every pixel pair in $G_A(G_B(y))$ and image y . The reason that we use this color loss calculation instead of an L2 distance in other color space is that the L2 metric only numerically measures the color difference, it cannot ensure that the color vectors have the same direction and the formulation is simple and fast for network computation.

3.4.4 Cyclic Perceptual Loss

Adversarial and cycle consistency losses are not able to preserve the textual and perceptual information of corrupted hazy images. Therefore, to achieve the perceptual quality we employed a cyclic perceptual loss. We utilized a pre-trained VGG16 model to extract features and calculated the distance between the features of hazy and reconstructed hazy images and also haze-free and the reconstructed haze-free counterpart using L2 norm (Simonyan and Zisserman, 2014). Equation 9 shows this loss function .

The goal of this loss function is to preserve the image structure and content features during dehazing and generate more realistic images. To calculate this loss, we focus on feature maps extracted from the 2nd and 5th pooling layers of VGG-16 pre-trained model. Equation 9 shows how this loss is calculated:

$$Loss_{CP} = \|(Vgg(G_B(G_A(x))) - Vgg(x))\|_2 + \|(Vgg(G_A(G_B(y))) - Vgg(y))\|_2 \quad (9)$$

where G_A and G_B are forward and backward generators, x belongs to domain X (*i.e.* the original domain, hazy images here) and y belongs to domain Y (*i.e.* the haze-free images). $G_B(G_A(x))$ and $G_A(G_B(y))$ are the reconstructed images. Vgg is a VGG16 feature extractor from the second and fifth pooling layers.

To calculate the L_{CP}^{Local} for the local discriminator we used the cropped local patches of input and output images and used the same equation 9.

4 EXPERIMENTS AND RESULTS

To evaluate the performance of our method compared to previous *paired and unpaired* methods, we train

a model on NYU dataset (Silberman et al., 2012) and test it on NYU dataset and also Middlebury dataset (Scharstein et al., 2014) as a cross-dataset to show how our model generalizes. NYU contains 1449 hazy images paired with their ground truth images and Middlebury contains 23 high-resolution(2k) hazy images with their ground truth. Since our method uses unpaired supervision, the training process received no information about which haze-free image corresponds to each hazy image.

4.1 Training

For training we need two sets of training datasets: trainA includes hazy images and trainB includes ground truth images (shuffled to simulate the unpaired supervision similar to other unpaired methods (Yang et al., 2018)). We opted for Adam optimizer (*momentum* = 0.5) with batch size of 1. Our initial learning rate was 0.0002 for the first 100 epochs, with linear decay to zero over the next 100 epochs. We implemented our model in PyTorch using two NVIDIA Tesla P100 GPUs and trained our network for 200 epochs.

4.2 Quality Measures

We used the following measurement metrics, to analyze the performance of our proposed method:

- **PSNR:** It measures the ratio between the maximum possible value of a signal and the power of distorting noise that affects the quality of its representation. The higher the PSNR, the more effective the reconstruction method is.
- **SSIM:** It is a Structural Similarity Index which is a perceptual metric that quantifies image quality degradation caused by processing. In this measurement, image degradation is considered as the change of perception in structural information (Kumar and Moyal, 2013).
- **CIEDE2000:** It measures the color difference between hazy and dehazed images; smaller values indicate better color preservation, thus better dehazing and perceptual quality (Luo et al., 2001).

4.3 Ablation Study

To demonstrate the effectiveness of the local discriminator, cyclic perceptual loss, and self-regularized color loss, we perform several ablation experiments.

Figure 3 depicts a couple of examples on how color loss helps with color artifacts removal. Employing color loss has enabled the network to remove artifacts effectively.

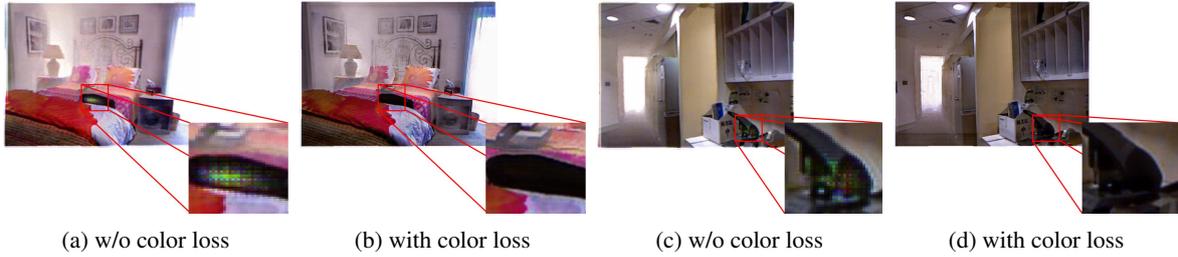


Figure 3: Examples showing the importance of color loss in our model ECDN.

PSNR: 12.14, SSIM: 0.74 **PSNR: 13.43, SSIM: 0.78** **PSNR: 16.10, SSIM: 0.74** **PSNR: 19.11, SSIM: 0.90**



Figure 4: Comparison between CycleGAN, Cycle-Dehaze and the proposed method.

Table 1: Ablation study over NYU dataset. The larger values of PSNR, SSIM and the smaller value of CIEDE2000 indicate better dehazing and perceptual quality.

Setting	↑ PSNR	↑SSIM	↓ CIEDE2000
CycleGAN	13.3879	0.5223	17.6113
ECDN w/o color loss	14.5402	0.7407	15.6401
ECDN w/o perceptual loss	14.6582	0.7312	15.6348
EDCN w/o residual blocks	14.1092	0.6923	16.4344
EDCN w/o local discriminator	14.0681	0.7111	19.9466
ECDN	16.0531	0.8244	14.9436

We compared our method with other cycle-consistent unpaired image-to-image translation methods. Figure 4 shows the comparison between CycleGAN, Cycle-dehaze and our method using an example image from NYU dataset. As one can observe our method removed more haze and the generated haze-free images is closer to the ground truth image. The red bounding boxes signify some parts of the image with different amount of haze removed by these methods.

Table 1 depicts the results of our ablation study in terms of PSNR, SSIM and CIEDE2000. One can observe that incorporating local discriminators can help achieve better PSNR, SSIM and CIEDE2000, meaning better restoration and generation of more visually pleasing results. The best results in terms of PSNR, SSIM, CIEDE2000 are achieved when the local discriminators, cyclic perceptual loss, and self-regularized color loss are incorporated.

4.4 Quantitative and Qualitative Analysis

We compare our model with both *paired* and *unpaired* methods, on the NYU and Middlebury datasets. Our method as well as the competitors are trained on the NYU dataset, and tested on NYU dataset and Middlebury dataset as a cross-dataset. Our method outperforms other methods in terms of SSIM and PSNR on both NYU and Middlebury datasets.

Table 2 and 3 and show the results on NYU and Middlebury datasets respectively. Our method outperforms the other methods in terms of SSIM and PSNR, and is the second best in terms of CIEDE2000.

Figure 5 shows the results of our method compared with other methods. DCP suffers from color distortion and over-exposure. CycleGAN introduces color artifacts and color shifting, and fails to remove much haze especially from dense hazy images. MSCNN and DehazeNet similarly fail to remove much haze from hazy images as well.

Table 2: Results on NYU dataset. Some of the numbers for the previous work are taken from (Yang et al., 2018; Engin et al., 2018).

Method	\uparrow PSNR	\uparrow SSIM	\downarrow CIEDE2000
DCP (He et al., 2010)	10.9803	0.6458	18.9781
CycleGAN (Cai et al., 2016)	13.3879	0.5223	17.6113
Cycle-Dehaze (Engin et al., 2018)	15.41	0.66	19.04432
DDN (Yang et al., 2018)	15.5456	0.7726	11.8414
DehazeNet (Cai et al., 2016)	12.8426	0.7175	15.8782
MSCNN (Ren et al., 2016)	12.2669	0.7000	17.4497
Ours	16.0531	0.8244	14.9436

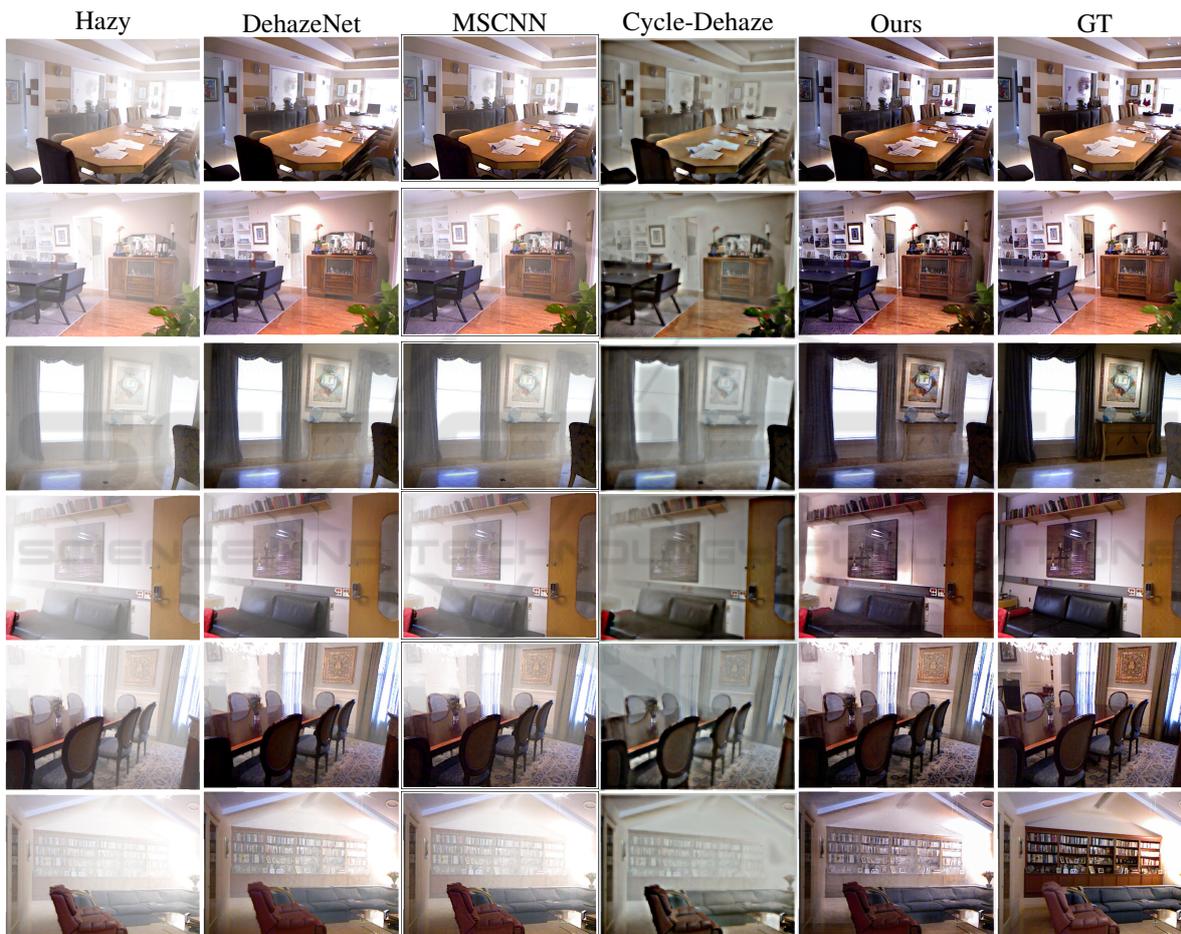


Figure 5: Comparison of the state-of-the-art dehazing methods on NYU dataset.

Our method, on the other hand is able to generate more natural haze-free images which are much closer to the ground truth image. Moreover, one can observe that our model outperforms the above-mentioned methods in recovery of details, and generates more natural images with least color artifacts.

5 CONCLUSION

In this paper, we treated the image dehazing problem as an image-to-image translation problem, and proposed a cycle-consistent generative adversarial network, called ECDN, for unpaired image dehazing. ECDN utilizes discriminators with a local-global structure and generators with an encoder-decoder ar-

Table 3: Results on Middlebury dataset. The numbers for the previous work are taken from (Yang et al., 2018; Engin et al., 2018).

Method	\uparrow PSNR	\uparrow SSIM
DCP (He et al., 2010)	12.0234	0.6902
CycleGAN (Cai et al., 2016)	11.3037	0.3367
Cycle-Dehaze (Engin et al., 2018)	15.6016	0.8532
DDN (Yang et al., 2018)	14.9539	0.7741
DehazeNet (Cai et al., 2016)	13.5959	0.7502
MSCNN (Ren et al., 2016)	13.5501	0.7365
Ours	15.8747	0.8601

chitecture with residual blocks and skip links to remove haze effectively. It also leverages different loss functions to generate realistic clean images. Using two benchmark test datasets, we showed the effectiveness of the proposed method. Our method outperforms other methods in terms of PSNR and SSIM.

ACKNOWLEDGMENTS

We would like to thank the VISAPP'21 anonymous reviewers for their valuable feedback. This work is partially supported by National Science Foundation grant IIS-1565328. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors, and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- Ancuti, C., Ancuti, C. O., De Vleeschouwer, C., and Bovik, A. C. (2016). Night-time dehazing by fusion. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 2256–2260. IEEE.
- Ancuti, C. O., Ancuti, C., Hermans, C., and Bekaert, P. (2010). 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.
- Anvari, Z. and Athitsos, V. (2019). A pipeline for automated face dataset creation from unlabeled images. In *Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments*, pages 227–235.
- Anvari, Z. and Athitsos, V. (2020). Evaluating single image dehazing methods under realistic sunlight haze. *arXiv preprint arXiv:2008.13377*.
- Cai, B., Xu, X., Jia, K., Qing, C., and Tao, D. (2016). Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 25(11):5187–5198.
- Chen, J., Chen, J., Chao, H., and Yang, M. (2018). Image blind denoising with generative adversarial network based noise modeling. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3155–3164.
- Emberton, S., Chittka, L., and Cavallaro, A. (2015). Hierarchical rank-based veiling light estimation for underwater dehazing.
- Engin, D., Genç, A., and Kemal Ekenel, H. (2018). Cycle-dehaze: Enhanced cyclegan for single image dehazing. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 825–833.
- He, K., Sun, J., and Tang, X. (2010). Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2341–2353.
- Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., Yang, J., Zhou, P., and Wang, Z. (2019). Enlighten-gan: Deep light enhancement without paired supervision. *arXiv preprint arXiv:1906.06972*.
- Kumar, R. and Moyal, V. (2013). Visual image quality assessment technique using fsim. *International Journal of Computer Applications Technology and Research*, 2(3):250–254.
- Kupyn, O., Budzan, V., Mykhailych, M., Mishkin, D., and Matas, J. (2018). Deblurgan: Blind motion deblurring using conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8183–8192.
- Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690.
- Li, B., Peng, X., Wang, Z., Xu, J., and Feng, D. (2017). Aod-net: All-in-one dehazing network. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 4770–4778.
- Lin, W.-A., Chen, J.-C., Castillo, C. D., and Chellappa, R. (2018). Deep density clustering of unconstrained faces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8128–8137.
- Lin, W.-A., Chen, J.-C., and Chellappa, R. (2017). A proximity-aware hierarchical clustering of faces. In *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, pages 294–301. IEEE.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., and Berg, A. C. (2016). Ssd: Single shot multibox detector. In *European conference on computer vision*, pages 21–37. Springer.
- Long, J., Shelhamer, E., and Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440.
- Luo, M. R., Cui, G., and Rigg, B. (2001). The development of the cie 2000 colour-difference formula: Ciede2000. *Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group*

- (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur, 26(5):340–350.
- McCartney, E. J. (1976). Optics of the atmosphere: scattering by molecules and particles. *New York, John Wiley and Sons, Inc., 1976. 421 p.*
- Meng, G., Wang, Y., Duan, J., Xiang, S., and Pan, C. (2013). Efficient image dehazing with boundary constraint and contextual regularization. In *Proceedings of the IEEE international conference on computer vision*, pages 617–624.
- Narasimhan, S. G. and Nayar, S. K. (2000). Chromatic framework for vision in bad weather. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662)*, volume 1, pages 598–605. IEEE.
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788.
- Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., and Yang, M.-H. (2016). Single image dehazing via multi-scale convolutional neural networks. In *European conference on computer vision*, pages 154–169. Springer.
- Scharstein, D., Hirschmüller, H., Kitajima, Y., Krathwohl, G., Nešić, N., Wang, X., and Westling, P. (2014). High-resolution stereo datasets with subpixel-accurate ground truth. In *German conference on pattern recognition*, pages 31–42. Springer.
- Silberman, N., Hoiem, D., Kohli, P., and Fergus, R. (2012). Indoor segmentation and support inference from rgb-d images. In *European conference on computer vision*, pages 746–760. Springer.
- Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Srinivasa, G. and Shree, K. (2002). Vision and the atmosphere. *International Journal of Computer Vision*, 48(3):233–254.
- Tan, R. T. (2008). Visibility in bad weather from a single image. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE.
- Tarel, J.-P. and Hautiere, N. (2009). Fast visibility restoration from a single color or gray level image. In *2009 IEEE 12th International Conference on Computer Vision*, pages 2201–2208. IEEE.
- Wang, R., Zhang, Q., Fu, C.-W., Shen, X., Zheng, W.-S., and Jia, J. (2019). Underexposed photo enhancement using deep illumination estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6849–6857.
- Yang, S., Luo, P., Loy, C.-C., and Tang, X. (2016). Wider face: A face detection benchmark. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5525–5533.
- Yang, X., Xu, Z., and Luo, J. (2018). Towards perceptual image dehazing by physics-based disentangle-ment and adversarial training. In *Thirty-second AAAI conference on artificial intelligence*.
- Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232.