

Article

Enhanced CycleGAN Network with Adaptive Dark Channel Prior for Unpaired Single-Image Dehazing

Yijun Xu ¹ , Hanzhi Zhang ², Fuliang He ^{2,3,*} , Jiachi Guo ¹ and Zichen Wang ¹

¹ Westa College, Southwest University, Chongqing 400715, China

² College of Electronic and Information Engineering, Southwest University, Chongqing 400715, China

³ Chongqing Key Laboratory of Nonlinear Circuits and Intelligent Information Processing, Southwest University, Chongqing 400715, China

* Correspondence: lighter@swu.edu.cn

Abstract: Unpaired single-image dehazing has become a challenging research hotspot due to its wide application in modern transportation, remote sensing, and intelligent surveillance, among other applications. Recently, CycleGAN-based approaches have been popularly adopted in single-image dehazing as the foundations of unpaired unsupervised training. However, there are still deficiencies with these approaches, such as obvious artificial recovery traces and the distortion of image processing results. This paper proposes a novel enhanced CycleGAN network with an adaptive dark channel prior for unpaired single-image dehazing. First, a Wave-Vit semantic segmentation model is utilized to achieve the adaption of the dark channel prior (DCP) to accurately recover the transmittance and atmospheric light. Then, the scattering coefficient derived from both physical calculations and random sampling means is utilized to optimize the rehazing process. Bridged by the atmospheric scattering model, the dehazing/rehazing cycle branches are successfully combined to form an enhanced CycleGAN framework. Finally, experiments are conducted on reference/no-reference datasets. The proposed model achieved an SSIM of 94.9% and a PSNR of 26.95 on the SOTS-outdoor dataset and obtained an SSIM of 84.71% and a PSNR of 22.72 on the O-HAZE dataset. The proposed model significantly outperforms typical existing algorithms in both objective quantitative evaluation and subjective visual effect.



Citation: Xu, Y.; Zhang, H.; He, F.; Guo, J.; Wang, Z. Enhanced CycleGAN Network with Adaptive Dark Channel Prior for Unpaired Single-Image Dehazing. *Entropy* **2023**, *25*, 856. <https://doi.org/10.3390/e25060856>

Academic Editors: Giuliana Ramella and Isabella Torricollo

Received: 12 April 2023

Revised: 22 May 2023

Accepted: 23 May 2023

Published: 26 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: dehaze; atmospheric scattering model; cycle generative adversarial network; dark channel prior

1. Introduction

With the rapid development of digital society, computer vision technology is increasingly being applied in the fields of autonomous driving, remote sensing imaging, and intelligent monitoring. However, the quality of the images acquired by photographic equipment in hazy weather is severely affected, with the target object being obscured and the image losing a lot of detailed information. Furthermore, degraded images are not conducive to subsequent high-level vision tasks. Therefore, a method for processing and clarifying hazy degraded images is highly desired.

At present, single-image dehazing has become a mainstream method for image clarification because it is cost-effective and requires no additional constraint information. Single-image dehazing methods can be classified into image enhancement, image restoration and learning-based approaches based on their mechanism.

Traditional image enhancement methods for dehazing include Retinex theory [1], histogram equalization [2], and wavelet transforms. These methods adjust the contrast and saturation of the image to achieve dehazing without considering the physical nature of haze formation. However, these global enhancements often cause the loss of some local information and perform poorly when facing hazy images in complex scenes. In recent

years, many methods combining image fusion have been widely proposed. Zheng et al. [3] used gamma correction to obtain a sequence of multiple exposed images from a single hazy image, then integrated the best region of saturation using adaptive decomposition to produce a clear image. Similarly, Galdran [4] utilized a multiscale Laplacian transform to fuse artificially exposed images to achieve dehazing. Zhu et al. [5] implemented feature extraction of a single image based on the idea of image space domain transformation; these authors then used a multiscale fusion algorithm based on fast filtering and saturation curve analysis to fuse the transformed image. These image fusion-based methods have improved traditional image enhancement, but the complexity of the algorithms used in these methods is high, and there are certain limitations to the stability of the dehazing they can provide.

Image restoration [6–9] uses prior knowledge or assumptions to establish a physical model of image degradation to achieve clarity. He et al. [7] discovered the dark channel prior (DCP) and mapped it to the atmospheric scattering model [10], designing an effective haze removal method. Zhu et al. [8] revealed the connection between haze concentration, image brightness, and saturation and developed the color attenuation prior. Berman et al. [9] proposed that haze alters the original tight color clusters in RGB space and forms haze lines through atmospheric light coordinates. Wang et al. [11] estimated the transmittance based on a prior for which there exists a linear relationship between the minimum channel of a hazy image and a clear image; these authors also introduced a weakening strategy combined with a quad-tree method of subdividing additional channels to restore the atmospheric light. Physical-model-based methods have been widely adopted since they are simple and efficient. Unfortunately, this type of algorithm requires high accuracy in parameter estimation and fails in regions that do not satisfy the prior; thus, the defogging results are often accompanied by negative effects such as color distortion and halos.

Recently, learning-based methods have significantly pushed the state of the art of unpaired image dehazing. Cai et al. [12] devised DehazeNet by integrating four different traditional defogging algorithms with deep learning. Zhang et al. [13] set two sub-networks in the pyramid network to obtain the transmittance and the atmospheric light, respectively. Li et al. [14] proposed a light-weight CNN network that they combined with the atmospheric scattering model to achieve dehazing. Chen et al. [15] developed GCANet on the basis of generative adversarial networks and adopted smooth convolution instead of extended convolution to solve the problem of grid artifacts. Similarly, a series of networks [16–19] were designed to derive clear images directly from the input hazy images without considering the degradation mechanism. Compared to conventional image enhancement and prior-based haze removal models, learning-based methods have achieved great progress. However, most of these methods are trained based on paired data and rely on clear images for positive supervision. This training process of supervision leads to excessive sensitivity to samples and the poor generalization of real-world haze removal. To address this issue, various unsupervised learning methods have been proposed. Li et al. [20] presented an unsupervised, unpaired defogging algorithm based on layer disentanglement, breaking away from training on large-scale datasets. However, this algorithm often produces images with serious color distortion and poor stability during defogging. Zhao et al. [21] proposed a weakly supervised RefinedNet, which combines the dark channel prior with a learning-based method using unpaired data for adversarial learning to improve the quality of the defogged images. Li et al. [22] integrated multi-scale feature representation with an attention mechanism and designed an enhanced decoder to improve the extraction of haze information. Ding et al. [23] unified the haze removal and noise suppression tasks and introduced a region similarity fusion module to obtain the final results. The development of unsupervised defogging algorithms has significantly alleviated the overfitting problem associated with supervised methods, but the defogging results lack realism, and the network structures tend to be complex, requiring higher computational resources.

Known as a powerful tool for unpaired image processing, CycleGAN (cycle generative adversarial network) [24] is characterized by its structure, which enables images to be

converted between two domains. Recently, many unpaired CycleGAN-based dehazing approaches have been widely proposed to solve the problem that paired samples are nearly unavailable in the real world. Engin et al. [25] designed a CycleDehaze system that combines a pyramid network for high-resolution images and introduces a cyclic perception loss to improve the dehazing quality. Zheng et al. [26] introduced an enhanced attention mechanism in the CycleGAN framework and applied it to the task of defogging remote sensing images. Most CycleGAN-based dehazing methods ignore the physical properties of the hazy environment; thus, the results lack realism and variability. In order to make progress on this issue, Yang et al. [27] combined CycleGAN with the atmospheric scattering model to recover the scene depth and haze density of images to improve dehazing quality and achieved better results on synthetic datasets; however, their network, with its high-complexity structure, is still limited regarding the accuracy of estimation for transmittance.

In this paper, we specifically propose a novel unpaired dehazing network termed ADCP-CycleGAN (adaptive DCP combined with CycleGAN). The network consists of two branches that implement the reconstruction of hazy and clear images, respectively. In the dehazing process, we use the scale-adaptive DCP to accurately recover the transmittance and atmospheric light and combine the variable scattering coefficient with depth to achieve a more realistic rehazing process.

The contributions of this paper can be summarized as follows:

- A novel unpaired single-image dehazing model is proposed to fuse the dark channel prior and the enhanced CycleGAN.
- An adaptive DCP is designed to rely on the Wave-ViT semantic segmentation model, and it can accurately recover the transmittance and atmospheric light.
- In the enhanced CycleGAN method, the scattering coefficient β is obtained from two different approaches in order to generate haze of various thicknesses and uneven distributions. β_1 is derived from the atmospheric scattering model, while β_2 is randomly sampled.

The article is organized as follows. Section 2 explains the preliminary knowledge of the atmospheric scattering model and the dark channel prior, as well as the basic structure of the cycle generative adversarial network. Section 3 elaborates the proposed image dehazing method based on CycleGAN with the adaptive dark channel. The experimental results, along with relevant discussions, are illustrated in Section 4. Conclusions and future work are summarized in Section 5.

2. Preliminaries

2.1. Atmospheric Scattering Model

To describe the mechanism of haze generation, McCartney et al. [10] proposed an atmospheric scattering model in 1977,

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

where $I(x)$ and $J(x)$ indicate a hazy degraded image and a clear image, respectively. A is the value of the global atmospheric light, and the transmission map, $t(x)$, can be derived from the following relationship:

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

where β is called the scattering coefficient, which can reflect the haze density. $d(x)$ is the depth of field.

Based on the atmospheric scattering model, a series of dehazing algorithms using prior knowledge [6–9] have been proposed. Among them, the most representative one is the dark channel prior [7] discovered by He et al.

2.2. Dark Channel Prior

The prior refers to certain pixels with lower intensities in at least one RGB channel as the dark channel, which can be represented as

$$J^{dark}(x) = \min_{y \in \Omega(x)} \left(\min_{c \in \{r,g,b\}} J^c(y) \right) \rightarrow 0, \quad (3)$$

where $J^c(y)$ denotes one of the RGB channels of a clear image, and $\Omega(x)$ is a patch centered on pixel x . The internal transmittance of $\Omega(x)$ can be approximated as a constant provided that the patch scale is sufficiently small. Substituting this into Equation (1), a mathematical derivation gives an estimate of the transmittance:

$$\tilde{t}(x) = 1 - \min_{y \in \Omega(x)} \left(\min_{c \in \{r,g,b\}} \frac{I^c(y)}{A^c} \right) \quad (4)$$

where $I^c(y)$ and A^c represent the original hazy image and the atmospheric ambient light in one of the RGB components, respectively. The subtraction term in Equation (4) is actually the dark channel intensity of $\frac{I^c(y)}{A^c}$. Combined with Equation (1), a clear result can be obtained as follows:

$$J(x) = \frac{I(x) - A}{\max(t(x), t_0)} + A, \quad (5)$$

in which t_0 is a tiny constant set to prevent the value of the denominator from being zero.

The patch size of the crucial parameter $\Omega(x)$ has a decisive impact on the defogging result. As shown in Figure 1b–d, an oversized patch ($\Omega(x) = 30$) would invalidate the assumption that “the transmittance in the patch is constant”, and the patch tends to cross the edge of the depth of field, leading to the halo effect. Conversely, as shown in Figure 1e–g, if the patch scale is too small ($\Omega(x) = 3$), the intensity of the dark pixels increases; thus, the transmittance obtained from Equation (4) is less than the real value, which may result in oversaturation, distortion, and an overall darkening of the image. Therefore, a single-scale $\Omega(x)$ will cause many unexpected negative effects and reduce the image quality.

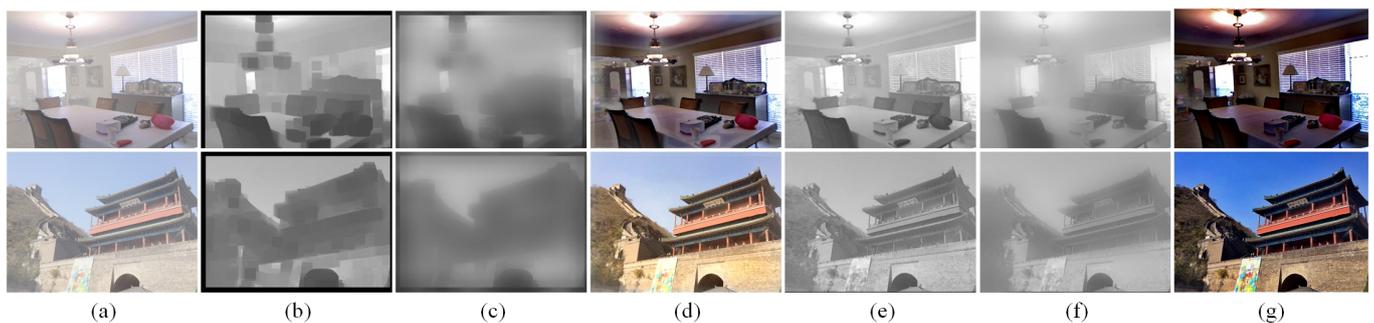


Figure 1. Effect of different patch sizes on dark channel prior (DCP) defogging. (a) Hazy input. (b–d) represent the dark channel map, transmission map, and dehazing result, respectively, based on $\Omega(x) = 30$. (e–g) are the corresponding groups, while $\Omega(x) = 3$.

Based on this, a number of algorithms were subsequently proposed to optimize DCP performance. Chen et al. [28] proposed the concept of a “bright channel”, as opposed to the dark channel, in order to solve the problem of the misalignment of brightness in dehazing results. Zhu et al. [29] and Jackson et al. [30] introduced the energy minimization theory and Raleigh scattering theory, respectively, to remove artifacts and halos. To some extent, these approaches that introduce external theories act as a correction and complement to the original DCP, while they also undermine the advantages of DCP, i.e., its efficiency and simplicity. From the perspective of parameter adaption, Song et al. [31] compared the defogging effect at different scales in detail and adaptively adjusted the scale range

of the dark channel according to the color and edge characteristics of the hazy image. Hu et al. [32] and Guo et al. [33] focused on segmenting the sky region, which does not satisfy the prior, to improve the accuracy of transmittance recovery. Inspired by previous research, we attempt to further subdivide the feature regions of the images and apply more accurate segmentation techniques to improve the quality of parameter adaptation. In Section 3.2, we will elaborate on the detailed optimization method.

2.3. CycleGAN

Cycle generative adversarial network was first designed by Zhu et al. [24]. It has a network structure with two generators and dual discriminators by mirror-symmetrizing the traditional GAN. Based on this special network structure, CycleGAN can convert images in the original and target domains without the supervision of paired datasets, a property that makes it widely preferred for unpaired dehazing tasks [25,26,34,35].

As shown in Figure 2, the previous CycleGAN-based dehazing networks contain a rehazing cycle and a dehazing cycle. In essence, most of them simply treat “hazy” and “clear” as two style domains for image transformation, with poor network interpretability and severe traces of artificial recovery. Specifically, the rehazing operation ignores real hazy environments that occur with various thicknesses and uneven distributions in the natural world, resulting in a large gap between the generated hazy images and the actual photographed hazy dataset. This means that the rehazing cycle has little significance for the enhancement of dehazing processing and can even negatively affect the quality of outputs, resulting in issues such as obvious artificial recovery traces and distortion.

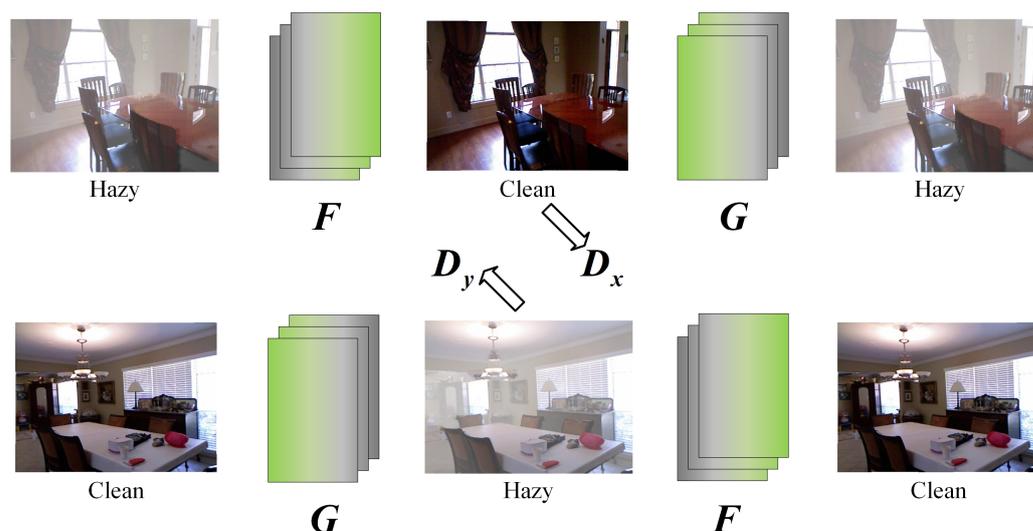


Figure 2. Structure of previous CycleGAN-based dehazing.

In order to improve the above issues, we introduce critical physical information to realize the enhancement of the dehazing and rehazing cycle. More details will be illustrated in Section 3.1.

3. Proposed Method

In this section, we elaborate on an unsupervised unpaired dehazing network termed ADCP-CycleGAN. We adopt adaptive DCP to accurately recover transmittance and atmospheric light for dehazing and achieve rehazing based on the depth and scattering coefficients. The two cycle branches of hazy/clarity reconstruction are connected by the atmospheric scattering model to form enhanced CycleGAN. The algorithm and network structure are detailed as follows.

3.1. Network Structure

The network consists of a hazy image reconstruction **H-H branch** and a clear image reconstruction **C-C branch**, as shown in Figure 3.

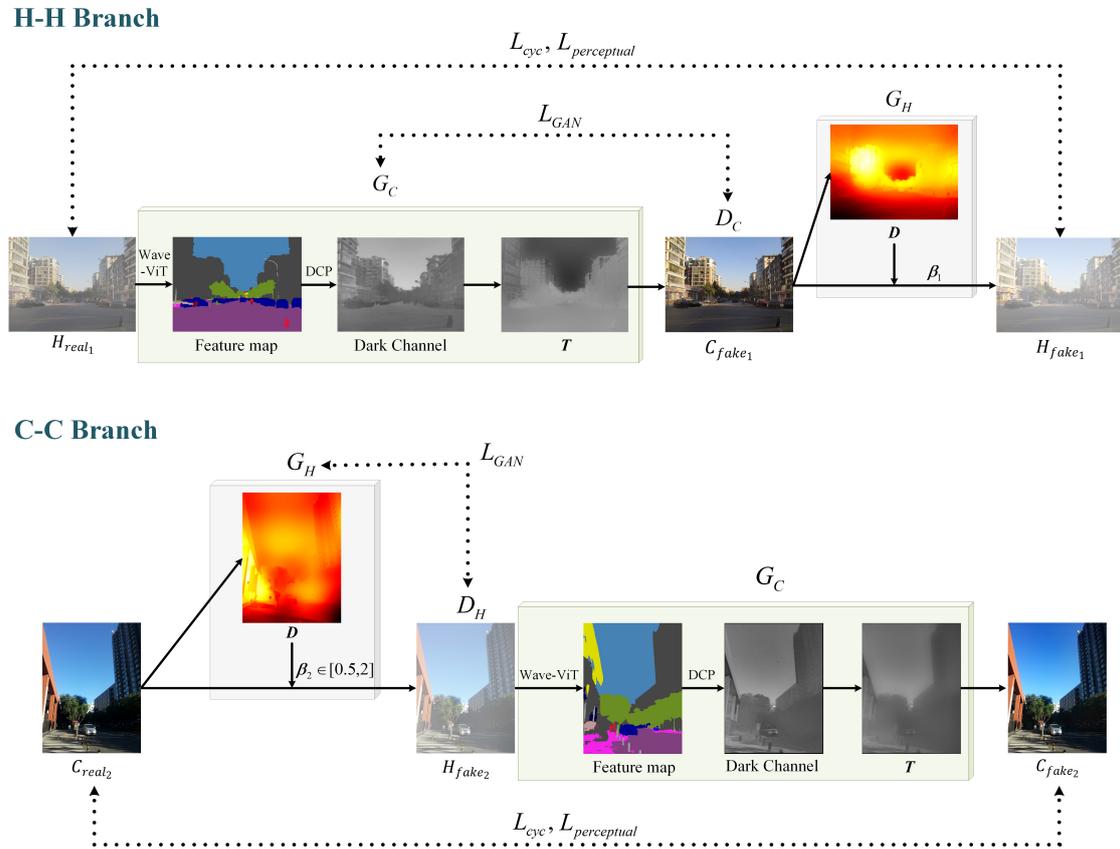


Figure 3. The structure of ADCP-CycleGAN.

H-H Branch. Given a hazy image H_{real_1} , we first perform Wave-ViT segmentation of the image to obtain the regional feature map. After the DCP operation, a dark channel map is obtained to deduce the transmittance T and atmospheric light A according to Equation (4). Then, the clear image C_{fake_1} can be acquired as follows:

$$C_{fake_1} = \frac{H_{real_1} - A}{T} + A \tag{6}$$

Based on the clear image, we can restore the depth D , at which time the scattering coefficient β_1 can be recovered to reflect the density of the haze distribution. With the depth and scattering coefficients, we ultimately obtain the reconstructed hazy H_{fake_1} . In this branch, the generator G_C is the dehazing processor, and D_C is the discriminator that identifies whether C_{fake_1} pertains to the clean domain.

C-C Branch. We initially derive the depth information from the input clear image C_{real_2} . The scattering coefficient β_2 is randomly sampled in the range of $[0.5, 2]$. The corresponding hazy image H_{fake_2} is subsequently acquired, and the same dehazing process as in the **H-H branch** is then adopted to acquire the final reconstructed clear image C_{fake_2} ; that is,

$$C_{fake_2} = \frac{H_{fake_2} - A}{T} + A \tag{7}$$

In this branch, the generator G_H produces the haze, and the discriminator D_H is used for recognizing if H_{fake_2} belongs to the hazy domain.

3.2. Adaptive DCP

In Section 2.2, we discussed in detail the drawbacks of the global fixedness of Ω in DCP. In this section, we continue the idea of parameter adaption to make further improvements.

To achieve a more refined segmentation of the feature regions, we here adopt the Wave-ViT model proposed by Yao et al. [36]. This model unites the wavelet transform with the Transformer network. With reversible downsampling for the lossless recovery of object texture details, it shows good performance in semantic segmentation tasks. The image division effect is shown in Figure 4.

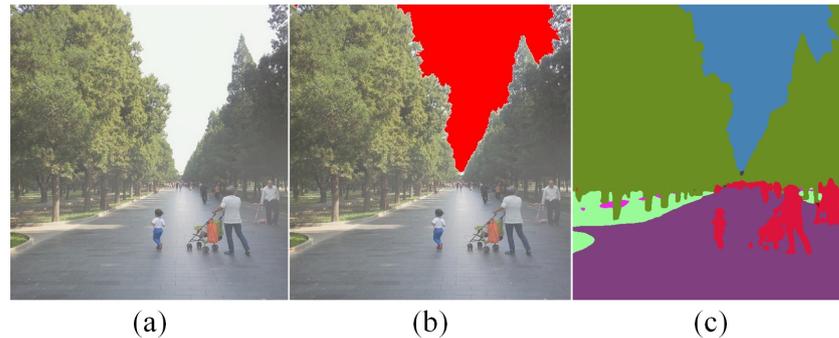


Figure 4. Hazy image region division. (a) Hazy image. (b) Sky region identification. (c) Foreground region segmentation.

We determine distinct patch sizes based on the essential properties of different areas in the image to achieve parameter self-adaption. The image can be divided into 3 regions: (a) The **Foreground** region, which consists of complex objects with rich colors and high saturation. An undersized patch may further aggravate the oversaturation phenomenon, whereas an oversized patch will violate the change of transmittance distribution in this region, causing an obviously distorted visual effect. Therefore, we set the patch scale of the foreground area in a normal interval that varies uniformly with the saturation. Specifically, the patch size of foreground area Ω_{fore} can be determined based on the saturation S and luminance L as follows:

$$L(x) = rI_r(x) + gI_g(x) + bI_b(x) \quad (8)$$

$$S(x) = 1 - \frac{\min_{c \in \{r,g,b\}} I_c(x)}{L(x)} \quad (9)$$

$$\Omega_{fore} = \max\{5, \text{round}(k \cdot \min_{c \in \{r,g,b\}} I_c(x))\} \quad (10)$$

where the max and round operators are used to set the patch scale as a positive integer. Based on previous research [7,28,31–33] on the defogging quality at different scales, we further conducted a validation experiment on the RESIDE dataset [37]. The results demonstrate that [5, 15] is a scale range that enables dark channel defogging to achieve optimal results, and other patch scales below or above this range experience significant negative effects, such as halos, luminance distortion, and oversaturation. Therefore, we take the value of k as 15 to ensure that Ω in the foreground region is adaptive in this range. As we calculate the brightness and saturation of the image in the HSI color space, the saturation value is in the range of [0, 1]. In order to change the patch scale uniformly with the saturation, we construct a linear mapping relationship between [5, 15] and [0, 1] so that pixel blocks with different levels of saturation in the foreground area can correspond to the suitable patches. (b) The **Sky** region has high brightness and low saturation. We choose a larger patch scale in the [25, 30] range in this area to intensify the defogging effect. At the same time, partitioning out the region helps us find the atmospheric light values using the

method in [7]. Notably, though we set the patch scale in the sky region to be much larger than in the foreground area, the sky area usually does not contain too much detail, the color saturation is more homogeneous, and the composition of the scene is simpler. At this point, the negative effects of large patches can be reduced. (c) The **Edge mutation** region. We set a smaller patch value in the range of [0, 3] in the depth of field border area to prevent halo effect and to preserve richer detail information.

3.3. Acquisition of Scattering Coefficient

In order to simulate the generation of real haze environments, which occur with various thicknesses and uneven distributions in the natural world, we optimize the rehazing process based on the atmospheric scattering model by combining the depth and density.

In the **H-H branch**, the scattering coefficient β_1 can be recovered according to Equation (2), as shown below:

$$\beta_1 = -\frac{\ln T}{D} \tag{11}$$

Based on this, the reconstructed hazy H_{fake_1} can be described as:

$$H_{fake_1} = C_{fake_1} e^{-\beta_1 D} + A(1 - e^{-\beta_1 D}) \tag{12}$$

Different from the **H-H branch**, the scattering coefficient β_2 in the **C-C branch** is randomly sampled in the range of [0.5, 2]. By altering the scattering coefficients, the generator G_H can produce hazy environments with arbitrary density distributions, as shown in Figure 5. Correspondingly, the hazy image H_{fake_2} is subsequently acquired as follows:

$$H_{fake_2} = C_{real_2} e^{-\beta_2 D} + A(1 - e^{-\beta_2 D}) \tag{13}$$

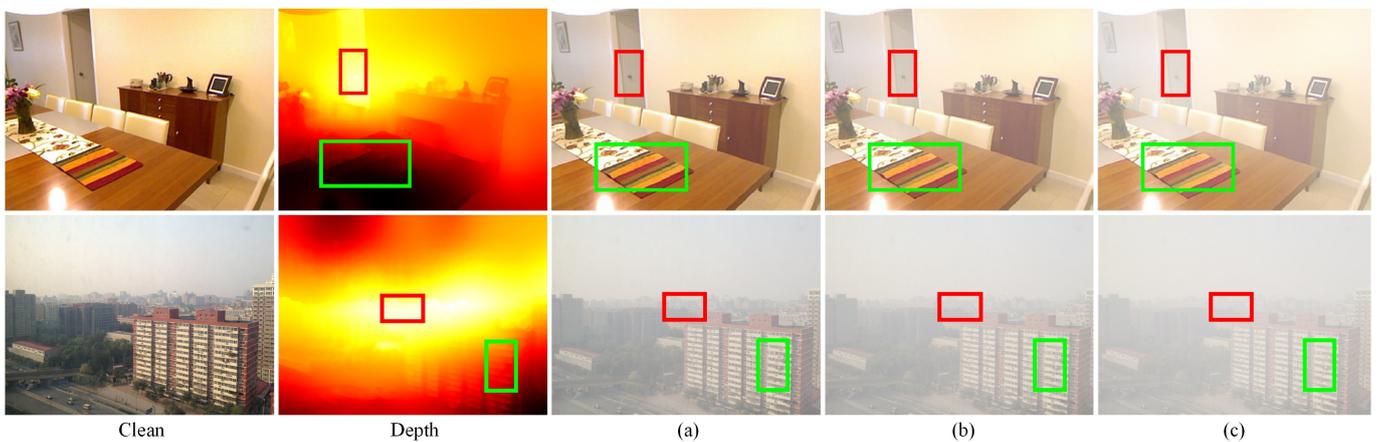


Figure 5. Different hazy images generated by G_H based on diverse depth-of-field and variable scattering coefficients. (a) $\beta = 0.5$. (b) $\beta = 1$. (c) $\beta = 2$.

It is noteworthy that, based on the atmospheric scattering model, the transmittance T and atmospheric light A derived from G_C can be applied in G_H to generate haze. Furthermore, these variable foggy images can also be used to augment the training of G_C . This mutually reinforcing haze removal/generation process constitutes the enhanced CycleGAN.

3.4. Calculation of Losses

GAN losses are incurred during the adversarial game between the generator and the discriminator. In our network, this occurs to ensure the quality of the dehazing and

rehazing process. In the **H-H branch**, the losses of the generator G_C and the discriminator D_C can be expressed as follows:

$$L_{GAN}(G_C) = E[(D_C(C_{fake_1}) - 1)^2] \quad (14)$$

$$L_{GAN}(D_C) = E[(D_C(C_{real_1}) - 1)^2] + E[(D_C(C_{fake_1}))^2] \quad (15)$$

in which C_{fake_1} is a clear image constructed by the generator G_C , and C_{real_1} is sampled from the clear image set $Set\{C\}$. In the **C-C Branch**, correspondingly, H_{fake_2} , which is derived from the rehazing generator G_H , and H_{real_2} , which is sampled from the hazy image $Set\{H\}$, are adopted to calculate the loss, which can be described as

$$L_{GAN}(G_H) = E[(D_H(H_{fake_2}) - 1)^2] \quad (16)$$

$$L_{GAN}(D_H) = E[(D_H(H_{real_2}) - 1)^2] + E[(D_H(H_{fake_2}))^2]. \quad (17)$$

Cycle-consistency losses calculate the consistency between the original and the target domain at both ends of the loop branch. In the **H-H branch**, the input H_{real_1} and the reconstructed hazy image H_{fake_1} must display sufficient levels of consistency. Likewise, C_{real_2} should agree with C_{fake_2} . Thus, the cycle-consistency losses can be written as Equation (18), where $\|\cdot\|_1$ denotes the L1 norm.

$$L_{cyc} = E_{H_{real_1} \sim Set\{H\}} \|H_{real_1} - H_{fake_1}\|_1 + E_{C_{real_2} \sim Set\{C\}} \|C_{real_2} - C_{fake_2}\|_1 \quad (18)$$

Cycle-perceptual losses. Although the cycle-consistency losses can be used to remove part of the noise, we also add cycle-perceptual losses to extract richer details and advanced features based on the VGG16 network to further enhance the structural similarity and ensure more realistic visual effects. The perceptual loss can be seen as Equation (19), where φ is the feature extractor and $\|\cdot\|_2$ denotes the L2 norm.

$$L_{perceptual} = \|\varphi(H_{real_1}) - \varphi(H_{fake_1})\|_2 + \|\varphi(C_{real_2}) - \varphi(C_{fake_2})\|_2 \quad (19)$$

Thus, the total loss function of ADCP-CycleGAN can be derived as:

$$L_{total} = \lambda_1 L_{GAN} + \lambda_2 L_{cyc} + \lambda_3 L_{perceptual} \quad (20)$$

λ_1 , λ_2 , and λ_3 are the weight-balancing factors of the three loss functions.

4. Experiment

4.1. Experimental Configuration

Datasets. In the experiment, four diverse datasets were adopted. (a) The RESIDE datasets [37] contain large amounts of hazy images synthesized artificially. SOTS-indoor and SOTS-outdoor contain 500 hazy/clear images indoors and outdoors, respectively, while ITS and OTS include 13,990 and 72,135 indoor and outdoor hazy and clear images, respectively. (b) The O-HAZE [38] dataset from the 2018 NTIRE Single Image Defogging Challenge contains 45 pairs of outdoor fogged/clear images with 10 pairs for testing. The images within this dataset are of high resolution and originate from real shots. (c) The BeDDE [39] dataset contains 208 real-world paired fogged/clear images of high quality captured in 23 different Chinese cities. We conducted qualitative comparison experiments on this dataset to evaluate its generalization ability and assess the subjective visual quality of real-world defogging effects. (d) In addition to the validation on the reference dataset, to compare the visual effects, we additionally introduced 30 hazy images captured in real life, as well as Fattal's dataset [40], which contains 31 real hazy images as non-reference samples.

Competitors & Metrics. We compared the proposed method with several state-of-art algorithms, including the most representative prior-based defogging algorithm, DCP [7]; supervised methods, including DehazeNet [12], GCANet [15], and FFANet [19]; and unsupervised methods, including ZID [20], RefineDNet [21], D4 [27], and USID [22]. For persuasive and reliable comparisons, the parameter settings were still implemented according to the content in Refs. [7,12,15,19–22,27]. We chose SSIM, PSNR [41], and LPIPS [42] as objective evaluation metrics for the dehazing performance on the reference dataset. For the test on the non-reference dataset, we focused on the evaluation of the visual effect of the dehazed image; thus, the information entropy (IE) and average gradient (AG) were employed to reflect the overall information and the local detail performance of the image, respectively. Moreover, we introduce the $NIQE$ [43] (natural image quality evaluator) metric, which can be expressed as

$$D(v_1, v_2, \Sigma_1, \Sigma_2) = \sqrt{(v_1 - v_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (v_1 - v_2)}, \quad (21)$$

where v_1, v_2, Σ_1 , and Σ_2 represent the mean MVG value and variance matrices of the natural and distorted image, respectively. $NIQE$ evaluates the test image by extracting features from the natural landscape, and its smaller value means the image is more compatible with human eye perception.

Training Settings. In the training phase, we randomly select 6000 images each from ITS and OTS, 380 images each from SOTS-indoor and SOTS-outdoor, and the training set from O-HAZE as input samples. Notably, due to the small sample size and high image resolution in the O-HAZE dataset, we cropped the 35 images into 700 copies to achieve sample expansion. All training images were rescaled to 256×256 . We set λ_1, λ_2 , and λ_3 as discussed in Section 3.4 to 0.2, 1, and 0.0001, respectively, to balance the weights of the three loss functions. The learning rate of the Adam optimizer was set to 0.0001, with a batch size of 2; furthermore, $\beta_1 = 0.5$, and $\beta_2 = 0.999$. We trained our model with an Nvidia GeForce RTX 2080 Ti graphics card and conducted our experiments on PyTorch.

4.2. Results on Reference and No-Reference Datasets

Comparison of reference datasets. Table 1 summarizes the average value of SSIM, PSNR, and LPIPS for every dehazing method tested on the SOTS-indoor (120 remaining images that differed from the training set, SOTS-outdoor (120 remaining images that differed from the training set), and O-HAZE datasets (10 test images cropped into 500 copies). On the SOTS-indoor test set, the supervised algorithms FFANet and GCANet demonstrate their strong capabilities and significant advantages. This is due to the fact that the supervised algorithms can sufficiently learn the image features based on paired datasets, thus performing well in simpler indoor scenes. Our proposed method achieves the best results among unsupervised algorithms. In outdoor haze removal, our algorithm performs the best among all nine algorithms on both SOTS-Outdoor and O-HAZE datasets, demonstrating that the proposed method maintains better generalization and high-quality defogging effects even in complex outdoor scenes. Meanwhile, it is worth noting that the supervised methods lose their dominant positions. To some extent, the results reflect the overfitting issues of supervised algorithms and their poor generalization abilities in handling complex scene defogging tasks.

Furthermore, we display visual comparisons in Figure 6. As can be observed, DCP results in an overall low brightness with obvious color distortion in the sky area. This is due to the fact that the prior is not met in the sky region. While ZID can remove haze, it suffers from significant color distortion in the fogged image. In the case of indoor defogging, RefineDNet produces some unpredictable noises in certain localized areas, such as the color block in the upper left corner of (g) and (h). The indoor defogging results of D4 suffer from serious over-brightening in the deep field due to its inaccurate estimation of atmospheric light, which is determined by taking the brightest pixel point as atmospheric light. This estimation method may lead to over-brightening of the image, especially in

indoor images with artificial noise, such as light sources and mirrors. On the other hand, FFANet, GCANet, and our supervised algorithm perform better in indoor defogging, with the former two being better at preserving the details of distant indoor objects. The outdoor defogging results, as shown in Figure 6c–f, reveal the overfitting problem of FFANet, as evidenced by the noticeable color halos on the gable roof in rows d and f. The proposed method exhibits better removal of residual haze in distant parts of the image, such as the distant buildings in row f. Overall, our algorithm shows good generalization ability for various types of defogging tasks, achieving thorough defogging and satisfactory subjective visual perception.

Table 1. Quantitative evaluation of nine algorithms on RESIDE and O-HAZE datasets. The best score is indicated in **Bolded**.

Type	Methods	SOTS-Indoor			SOTS-Outdoor			O-HAZE		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Prior	DCP [7]	16.61	0.855	0.225	19.41	0.861	0.122	12.32	0.516	0.473
Supervised	DehazeNet [12]	19.82	0.821	0.186	24.75	0.927	0.065	16.47	0.624	0.229
	GCANet [15]	30.23	0.975	0.161	24.36	0.894	0.115	18.51	0.693	0.332
	FFANet [19]	34.31	0.977	0.152	21.23	0.835	0.173	18.36	0.829	0.170
Unsupervised	RefineDNet [21]	25.06	0.929	0.199	23.58	0.914	0.047	19.27	0.853	0.152
	ZID [20]	17.26	0.801	0.244	12.19	0.614	0.396	9.82	0.437	0.528
	D4 [27]	25.40	0.934	0.207	25.75	0.936	0.035	19.90	0.844	0.147
	USID [22]	20.09	0.873	0.218	24.97	0.930	0.044	20.12	0.862	0.140
	Ours	25.98	0.941	0.157	26.95	0.949	0.031	22.72	0.871	0.136

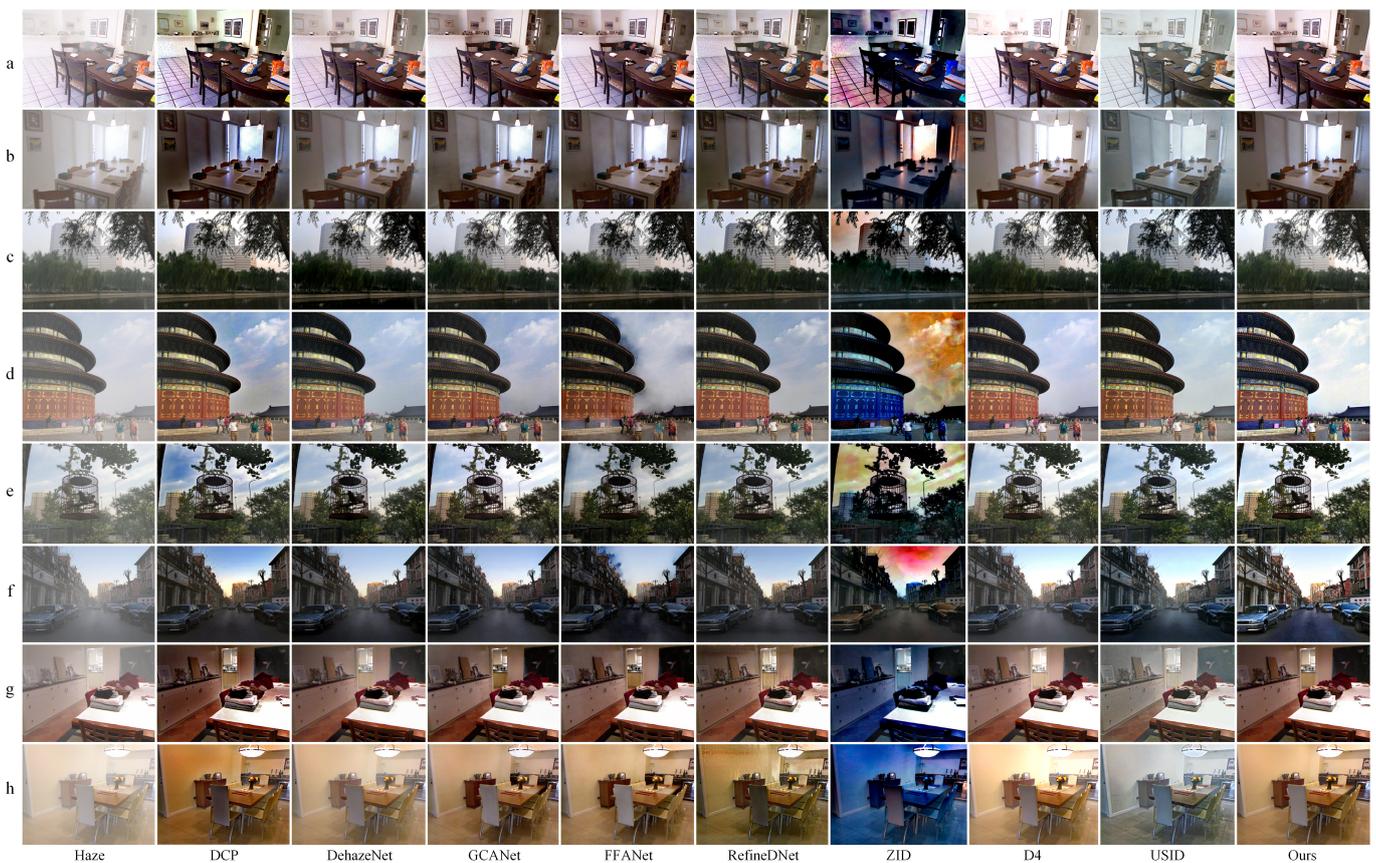


Figure 6. Comparative test of nine algorithms on RESIDE datasets. Rows (a),(b),(g),(h) show indoor defogging results, and rows (c–f) show outdoor defogging results. ADCP-CycleGAN dehazes well in different defogging scenarios.

Additionally, we compared the number of trainable parameters and the running time of our proposed ADCP-CycleGAN with other methods under the same experimental environment and summarized the results in Table 2. Of these methods, the prior-based DCP [7] does not require trainable parameters, and USID [22] outperforms the other algorithms in terms of the number of parameters and running time since it does not rely on calculating physical parameters in the atmospheric scattering model. The proposed method demonstrates acceptable network complexity and defogging efficiency, with fewer parameters and faster running speed compared to other state-of-the-art algorithms.

Table 2. Comparison of the number of trainable parameters and average running time of different dehazing methods.

Type	Methods	Number of Parameters	Runtime (s)
Prior	DCP [7]	-	0.2930
Supervised	DehazeNet [12]	0.008×10^6	1.6200
	GCANet [15]	0.660×10^6	0.9275
	FFANet [19]	4.964×10^6	1.3418
Unsupervised	RefineDNet [21]	63.378×10^6	0.7053
	ZID [20]	48.232×10^6	57.3681
	D4 [27]	11.707×10^6	0.0579
	USID [22]	4.022×10^6	0.0432
	Ours	4.275×10^6	0.0656

Comparison on no-reference real datasets. To verify the generalization ability of the network and the realism of the defogging results, we additionally introduced no-reference datasets. The quantitative evaluation results are reported in Table 3. Our method obtains the best scores in all three metrics, which means that the defogged images achieve acceptable results in terms of information content, detail representation, and visual effect.

Table 3. Quantitative evaluation of nine algorithms on no-reference datasets. The best score is indicated in **Bolded**.

Type	Methods	IE \uparrow	AG \uparrow	NIQE \downarrow
Prior	DCP [7]	7.2658	7.4342	9.3858
Supervised	DehazeNet [12]	7.2945	7.0627	7.7306
	GCANet [15]	7.3098	6.4246	6.9544
	FFANet [19]	7.1056	7.1901	7.3398
Unsupervised	RefineDNet [21]	7.0903	7.9750	6.8427
	ZID [20]	7.2770	5.1849	12.4221
	D4 [27]	7.2251	7.4858	7.1425
	USID [22]	7.3560	8.0217	7.0951
	Ours	7.5238	9.8605	6.5364

In order to demonstrate the defogging effect more clearly, we framed some local details of the image and zoomed in for comparison, as shown in Figure 7. In rows a and c, we framed the text area and zoomed in. Our method successfully preserved more edge details and restored the text information well. For the nature landscape image in row b, our method has effectively removed the residual haze, resulting in a natural color perception of the defogged image. Though GCANet also produces results with less residual haze, there are noticeable distortions in the sky region in rows b and d. FFANet defogging in real-world scenes is not desirable, as there are noticeable haze residues in the results and a large number of artifacts in the sky area of row d. In addition, the hues of USID dehazing results in rows c and d lacking naturalness and realism, resulting in poor visual effects. To

summarize, our algorithm consistently shows good defogging performance on real-world no-reference datasets, providing appealing subjective and visually realistic effects.

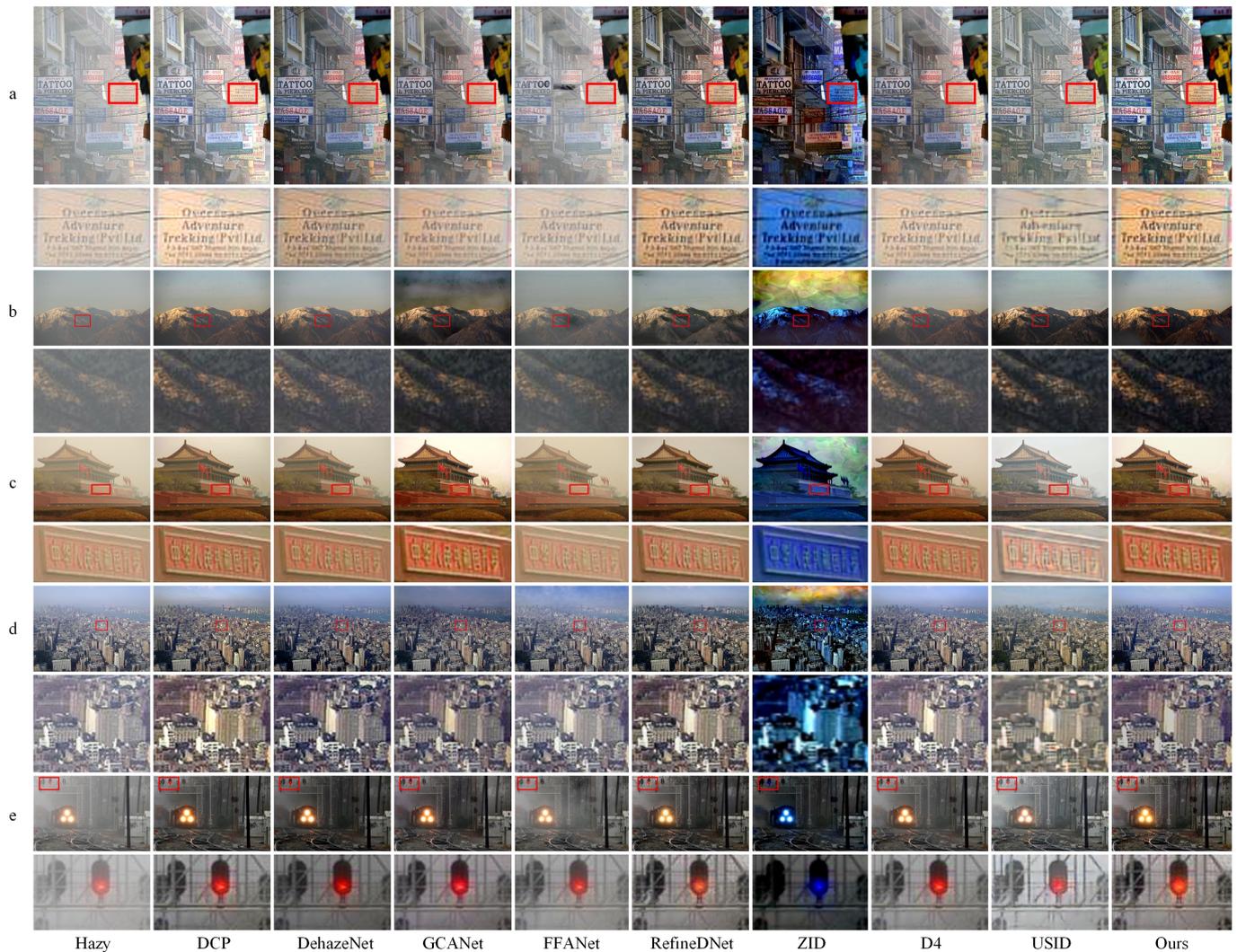


Figure 7. Comparative test of nine algorithms on no-reference datasets. The proposed method shows strong generalization ability and robustness in various real-world defogging tasks, with satisfactory overall picture quality and detailed information performance (a–e).

In addition, we conducted abundant extended experiments on the BeDDE dataset, and the visual comparison of the defogging results is shown in Figures 8 and 9. The satisfactory defogging effects further reveal the strong generalization ability and defogging stability of the proposed algorithm.

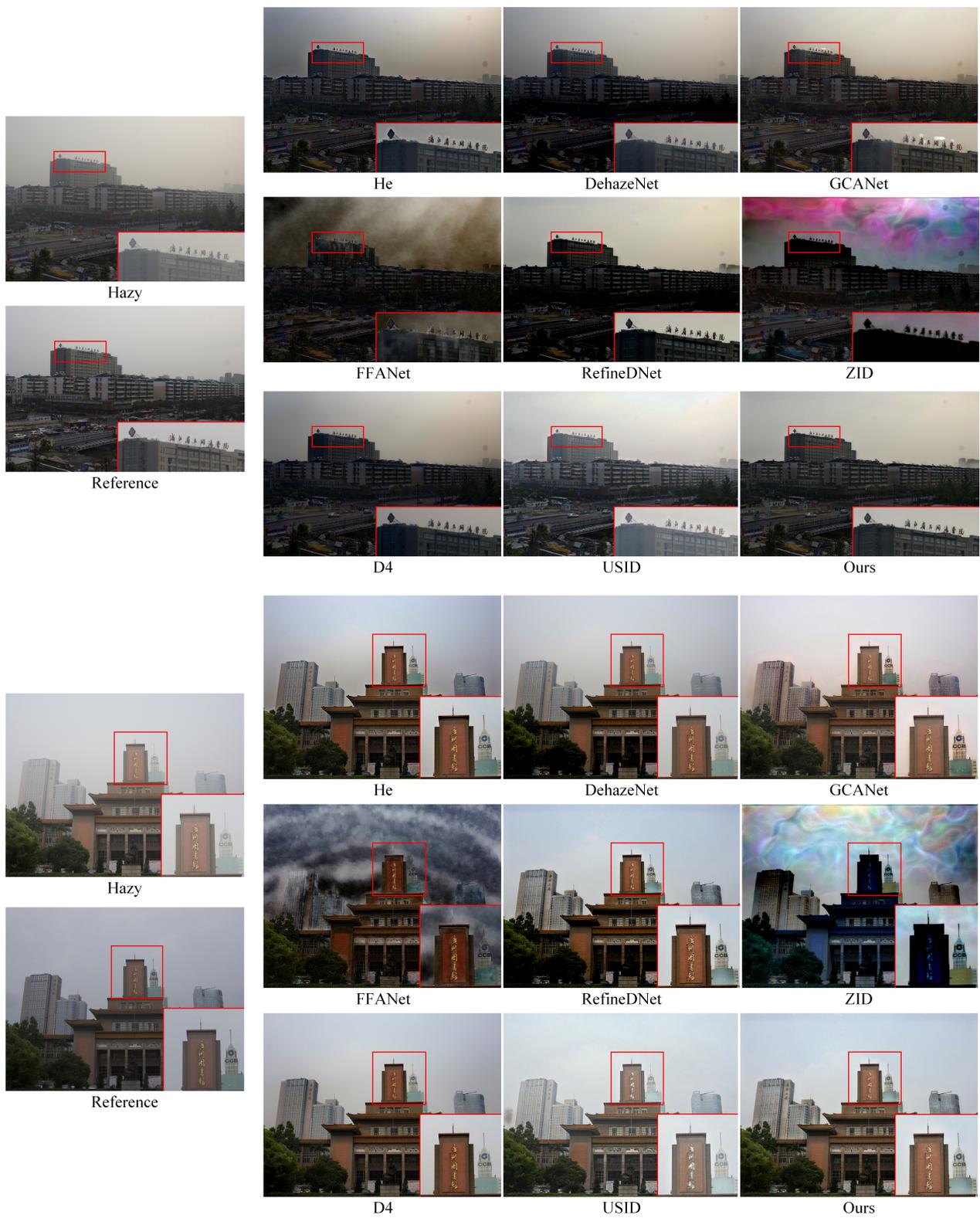


Figure 8. Comparative test of nine algorithms on BeDDE datasets.



Figure 9. Comparative test of nine algorithms on BeDDE datasets.

4.3. Ablation Study

To verify the effectiveness of the different components in ADCP-CycleGAN, we conducted an ablation study on the network. Three additional models were trained and compared with our proposed model on the SOTS dataset as follows: (a) Model A removes the Wave-ViT semantic segmentation and parameter adaptation module, and thus, the transmittance and the atmospheric light are recovered by the original DCP method; (b) the value of the scattering coefficient in Model B is set to a fixed constant; and (c) Model C deletes the cycle perceptual loss.

The dehazing results of the four models are reported in Figure 10. After removing the semantic segmentation module for the parameter adaptation of DCP, the defogging results of Model A show an obvious distortion. For the areas where the prior fails (such as the white floor tiles in row b), the distortion phenomenon appears, and the brightness of the picture in row c is also significantly darker. Model B lacks realism in the subjective visual effect of the defogging result since the scattering coefficient is set to a fixed value. Model C has a degraded performance regarding detail recovery after removing the cycle perceptual loss. The flowers in the far field in row c show an oversaturation of color, indicating that the deletion of cycle perceptual loss has an impact on defogging stability.

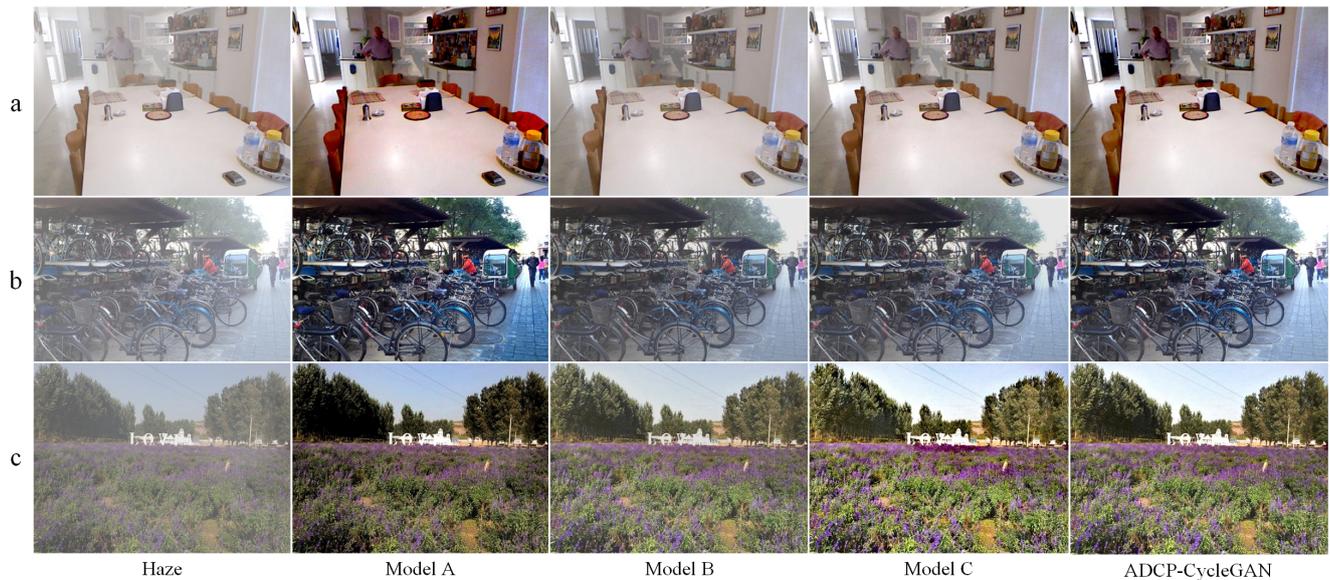


Figure 10. Ablation study on RESIDE datasets. The defogging results of Model A showed obvious oversaturation and color distortion. Model B showed significant haze residue. The defogging stability of Model C was reduced compared with ADCP-CycleGAN (a–c).

It is worth noting that in the quantitative analysis (as shown in Table 4), the degradation of Model B compared to ADCP-CycleGAN is different in indoor and outdoor dehazing tasks. This may be due to the fact that the haze distribution in indoor scenes is more uniform compared to outdoors; thus, the scattering coefficient may have a more significant impact on the outdoor haze removal. This also confirms that the scattering coefficient is not negligible in the outdoor dehazing.

Table 4. Quantitative evaluation of ablation study. The best score is indicated in **Bolded**.

Methods	SOTS-Indoor			SOTS-Outdoor		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Model A	21.60	0.872	0.209	22.78	0.904	0.046
Model B	24.22	0.921	0.166	24.19	0.916	0.043
Model C	23.09	0.919	0.173	24.48	0.925	0.037
ADCP-CycleGAN	25.98	0.941	0.157	26.95	0.949	0.031

5. Conclusions and Future Work

In this paper, we propose ADCP-CycleGAN, a novel enhanced CycleGAN network with adaptive DCP for unpaired single-image dehazing. In the network, we achieve the parameter adaption of DCP through a Wave-ViT semantic segmentation model to recover the transmittance and atmospheric light accurately. We optimize the rehazing process by deriving the scattering coefficient from both physical calculation and random sampling means to simulate the real haze distribution. The atmospheric scattering model is applied

to realize the connection between the dehazing and rehazing branch in order to build the enhanced CycleGAN. The extended experiments on both reference/no-reference datasets with diverse evaluation metrics confirm the effectiveness of our method. Specifically, our approach can generate haze that is more consistent with real-world scenarios based on depth and density. This could be particularly meaningful for tasks that require clear vision but lack unpaired datasets, such as remote sensing images, autonomous driving, and intelligent monitoring. Furthermore, we hope that our innovative combination of physical prior models with CycleGAN for dehazing can contribute to future developments in unsupervised learning for low-level vision tasks. However, there are also some aspects of our algorithm that deserve improvement. The accuracy of the depth estimation of the proposed method is affected when there is noise such as strong light and obscuration in the image. Meanwhile, due to the incorporation of a physical model in the proposed method, its inherent limitations may result in the local over-enhancement in a few defogged results. In our future work, we will also investigate the post-processing of the defogged images to further improve the image quality.

Author Contributions: Methodology, Y.X.; Software, H.Z.; Supervision, Project administration, F.H.; Data curation, J.G.; Writing—review and editing, Z.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the National Training Program of Innovation and Entrepreneurship for Undergraduates (Grant No. 202310635096), Venture Innovation Support Program for Chongqing Overseas Returnees (Grant No. cx2019133) and the Chongqing Research Project of the Foal Eagle Program (Grant No. CY220231).

Data Availability Statement: Data sharing no applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jobson, D.; Rahman, Z.; Woodell, G. A multi-scale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Trans. Image Process.* **1997**, *6*, 965–976. [[CrossRef](#)]
2. Zou, X.; Liu, Y.; Tan, Z. A fog-removing treatment based on combining high-frequency emphasis filtering and histogram equalization. *Key Eng. Mater.* **2011**, *474*, 2198–2202. [[CrossRef](#)]
3. Zheng, M.; Qi, G.; Zhu, Z.; Li, Y.; Wei, H.; Liu, Y. Image dehazing by an artificial image fusion method based on adaptive structure decomposition. *IEEE Sens. J.* **2020**, *20*, 8062–8072. [[CrossRef](#)]
4. Galdran, A. Image dehazing by artificial multiple-exposure image fusion. *Signal Process.* **2018**, *149*, 135–147. [[CrossRef](#)]
5. Zhu, Z.; Wei, H.; Hu, G.; Li, Y.; Qi, G.; Mazur, N. A novel fast single image dehazing algorithm based on artificial multiexposure image fusion. *IEEE Trans. Instrum. Meas.* **2020**, *70*, 1–23. [[CrossRef](#)]
6. Fattal, R. Single Image Dehazing. *ACM Trans. Graph.* **2008**, *27*, 1–9. [[CrossRef](#)]
7. He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *33*, 2341–2353.
8. Zhu, Q.; Mai, J.; Shao, L. A fast single image haze removal algorithm using color attenuation prior. *IEEE Trans. Image Process.* **2015**, *24*, 3522–3533.
9. Berman, D. Non-local image dehazing. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1674–1682.
10. McCartney, E.J. Optics of the atmosphere: Scattering by molecules and particles. *Phys. Bull.* **1977**, *28*, 521. [[CrossRef](#)]
11. Wang, W.; Yuan, X.; Wu, X.; Liu, Y. Fast image dehazing method based on linear transformation. *IEEE Trans. Multimed.* **2017**, *19*, 1142–1155. [[CrossRef](#)]
12. Cai, L.B.; Xu, X.M.; Jia, K.; Qing, C.M.; Tao, D.C. Dehazenet: An end-to-end system for single image haze removal. *IEEE Trans. Image Process.* **2016**, *25*, 5187–5198. [[CrossRef](#)] [[PubMed](#)]
13. Zhang, H.; Patel, V.M. Densely connected pyramid dehazing network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 3194–3203.
14. Li, B.; Peng, X.; Wang, Z.; Xu, J.; Feng, D. Aod-net: All-in-one dehazing network. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 4770–4778.
15. Chen, D.; He, M.; FAN, A. Gated context aggregation network for image dehazing and deraining. In Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 7–11 January 2019; pp. 1375–1383.
16. Li, R.; Pan, J.; Li, Z.; Tang, J. Single image dehazing via conditional generative adversarial network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8202–8211.
17. Qu, Y.; Chen, Y.; Huang, J.; Xie, Y. Enhanced pix2pix dehazing network. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 8160–8168.

18. Liu, X.; Ma, Y.; Shi, Z.; Chen, J. Griddehazenet: Attention-based multi-scale network for image dehazing. In Proceedings of the IEEE International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 7314–7323.
19. Qin, X.; Wang, Z.; Bai, Y.; Xie, X.; Jia, H. FFA-Net: Feature fusion attention network for single image dehazing. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 11908–11915.
20. Li, B.; Gou, Y.; Liu, J.Z.; Zhu, H.; Zhou, J.T.; Peng, X. Zero-shot image dehazing. *IEEE Trans. Image Process.* **2020**, *29*, 8457–8466. [[CrossRef](#)] [[PubMed](#)]
21. Zhao, S.; Zhang, L.; Shen, Y.; Zhou, Y. RefinedNet: A weakly supervised refinement framework for single image dehazing. *IEEE Trans. Image Process.* **2021**, *30*, 3391–3404. [[CrossRef](#)]
22. Li, J.; Li, Y.; Zhuo, L.; Kuang, L.; Yu, T. USID-Net: Unsupervised single image dehazing network via disentangled representations. *IEEE Trans. Multimed.* **2022**. [[CrossRef](#)]
23. Ding, B.; Zhang, R.; Xu, L.; Liu, G.; Yang, S.; Liu, Y.; Zhang, Q. U^2D^2 Net: Unsupervised Unified Image Dehazing and Denoising Network for Single Hazy Image Enhancement. *IEEE Trans. Multimed.* **2023**. [[CrossRef](#)]
24. Zhu, J.; Park, T.; Isola, P. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.
25. Engin, D.; Genc, A.; Ekenel, H.K. Cycledehaze: Enhanced cycleGAN for single image dehazing. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–23 June 2018; pp. 825–833.
26. Zheng, Y. Dehaze-AGGAN: Unpaired Remote Sensing Image Dehazing Using Enhanced Attention-Guide Generative Adversarial Networks. *IEEE Trans. Geosci. Remote. Sens.* **2022**, *60*, 1–13. [[CrossRef](#)]
27. Yang, Y.; Wang, C.; Liu, R.; Zhang, L.; Guo, X.; Tao, D. Self-augmented Unpaired Image Dehazing via Density and Depth Decomposition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 2037–2046.
28. Chen, Y.; Lu, C.T. Single Image Dehazing Based on Superpixel Segmentation Combined with Dark-Bright Channels. *Laser Optoelectron. Prog.* **2020**, *57*, 161023. [[CrossRef](#)]
29. Zhu, M.; He, B.; Wu, Q. Single image dehazing based on dark channel prior and energy minimization. *IEEE Signal Process. Lett.* **2017**, *25*, 174–178. [[CrossRef](#)]
30. Jackson, J.; Kun, S.; Agyekum, K.O. A fast single-image dehazing algorithm based on dark channel prior and rayleigh scattering. *IEEE Access* **2020**, *8*, 73330–73339. [[CrossRef](#)]
31. Song, Y.; Luo, H.; Hui, B.; Chang, Z. Haze removal using scale adaptive dark channel prior. *Infrared Laser Eng.* **2016**, *45*, 928002. [[CrossRef](#)]
32. Hu, Q.; Zhang, Y.; Zhu, Y.; Jiang, Y.; Song, M. Single image dehazing algorithm based on sky segmentation and optimal transmission maps. *Vis. Comput.* **2023**, *39*, 997–1013. [[CrossRef](#)]
33. Guo, F.; Qiu, J.; Tang, J. Single Image Dehazing Using Adaptive Sky Segmentation. *IEEJ Trans. Electr. Electron. Eng.* **2021**, *16*, 1209–1220. [[CrossRef](#)]
34. Liu, W.; Hou, X.; Duan, J.; Qiu, G. End-to-end single image fog removal using enhanced cycle consistent adversarial networks. *IEEE Trans. Image Process.* **2020**, *29*, 7819–7833. [[CrossRef](#)]
35. Zhao, J.; Zhang, J.; Li, Z.; Hwang, J.N.; Gao, Y.; Fang, Z.; Jiang, X.; Huang, B. Dd-cycleGAN: Unpaired image dehazing via double-discriminator cycle-consistent generative adversarial network. *Eng. Appl. Artif. Intell.* **2019**, *82*, 263–271. [[CrossRef](#)]
36. Yao, T.; Pan, Y.; Li, Y.; Ngo, C.W.; Mei, T. Wave-vit: Unifying wavelet and transformers for visual representation learning. In Proceedings of the Computer Vision-ECCV 2022: 17th European Conference, Tel Aviv, Israel, 23–27 October 2022; pp. 328–345.
37. Li, B.; Ren, W.; Fu, D.; Tao, D.; Feng, D.; Zeng, W.; Wang, Z. Benchmarking singleimage dehazing and beyond. *IEEE Trans. Image Process.* **2018**, *28*, 492–505. [[CrossRef](#)]
38. Ancuti, C.O.; Ancuti, C.; Timofte, R.; Vleeschouwer, C.D. O-HAZE: A Dehazing Benchmark with Real Hazy and Haze-Free Outdoor Images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 754–762.
39. Zhao, S.; Zhang, L.; Huang, S.; Shen, Y.; Zhao, S.; Yang, Y. Evaluation of defogging: A real-world benchmark dataset, a new criterion and baselines. In Proceedings of the 2019 IEEE International Conference on Multimedia and Expo (ICME), Shanghai, China, 8–12 July 2019; pp. 1840–1845.
40. Fattal, R. Dehazing using color-lines. *ACM Trans. Graph.* **2014**, *34*, 1–14. [[CrossRef](#)]
41. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)]
42. Zhang, R.; Isola, P.; Efros, A.A.; Shechtman, E.; Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 586–595.
43. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a “completely blind” image quality analyzer. *IEEE Signal Process. Lett.* **2012**, *20*, 209–212. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.