



## Article

# A Lightweight Forest Scene Image Dehazing Network Based on Joint Image Priors

Xixuan Zhao <sup>1,2,3</sup> , Yu Miao <sup>1,2</sup>, Zihui Jin <sup>1,2</sup>, Jiaming Zhang <sup>1,2</sup>  and Jiangming Kan <sup>1,2,3,\*</sup>

<sup>1</sup> School of Technology, Beijing Forestry University, No. 35 Tsinghua East Road, Haidian District, Beijing 100083, China; zhaoxixuan@bjfu.edu.cn (X.Z.)

<sup>2</sup> Key Laboratory of State Forestry Administration on Forestry Equipment and Automation, No. 35 Tsinghua East Road, Haidian District, Beijing 100083, China

<sup>3</sup> Foshan-Zhongke Innovation Research Institute of Intelligent Agriculture and Robotics, Jingu Zhichuang Industrial Community, No. 2 Yongan North Road, Dawei Community, Guicheng Street, Nanhai District, Foshan 528251, China

\* Correspondence: kanjm@bjfu.edu.cn

**Abstract:** Fog interference is an unfavorable issue when using vision sensors to monitor forest environmental resources. The existence of fog causes intelligent forest vision sensor equipment to fail to obtain accurate information on environmental resources. Therefore, this study proposes a lightweight forest scene image dehazing network to remove fog interference from the vision system. To deal with the extraction of detailed forest image features, we propose utilizing joint image priors including white balance, contrast, and gamma correction feature maps as inputs of the network to strengthen the learning ability of the deep network. Focusing on reducing the computational cost of the network, four different kinds of Ghost Bottleneck blocks, which adopt an SE attention mechanism to better learn the abundant forest image features for our network, are adopted. Moreover, a lightweight upsampling module combining a bilinear interpolation method and a convolution operation is proposed, thus reducing the computing space used by the fog removal module in the intelligent equipment. In order to adapt to the unique color and texture features of forest scene images, the cost function consisting of  $L_1$  loss and multi-scale structural similarity (MS-SSIM) loss is specially designed to train the proposed network. The experimental results show that our proposed method obtains more natural visual effects and better evaluation indices. The proposed network is trained both on indoor and outdoor synthetic datasets and tested on synthetic and real foggy images. The PSNR achieves an average value of 26.00 dB and SSIM achieves 0.96 on the indoor synthetic dataset, while PSNR achieves an average value of 25.58 dB and SSIM achieves 0.94 on the outdoor synthetic test images. The average processing time of our proposed dehazing network for a single foggy image with a size of  $480 \times 640$  is 0.26 s.

**Keywords:** forestry scene; image dehazing; prior knowledge; lightweight network



**Citation:** Zhao, X.; Miao, Y.; Jin, Z.; Zhang, J.; Kan, J. A Lightweight Forest Scene Image Dehazing Network Based on Joint Image Priors. *Forests* **2023**, *14*, 2062. <https://doi.org/10.3390/f14102062>

Academic Editors: Faramarz F. Samavati and Ali Mahdavi-Amiri

Received: 11 September 2023

Revised: 8 October 2023

Accepted: 13 October 2023

Published: 16 October 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/>).

## 1. Introduction

Fog interference is an unfavorable issue when using vision sensors to monitor forest environmental resources. Forest areas have unique climatic conditions with large temperature changes, abundant heat, and precipitation, leading to easily generated fog all year round, thus resulting in the degradation of image quality acquired via vision sensors of intelligent forestry equipment, which is like covered eyes. Degraded images present negative features such as texture loss, low contrast, color distortion, and dynamic range compression [1]. Obtaining inaccurate information about forest resources results in greatly reduced identification accuracy of forest fruits and standing trees, as well as highly increased false detection rates via monitoring systems, which impairs the effectiveness and availability of information obtained via vision sensors in forests. In addition, forest scenes have unique prior characteristics, such as uneven light distribution, single-color features,

and rich texture details, which makes the accurate operation of forestry vision systems more challenging. Therefore, combining unique image priors to solve the forest scene image dehazing problem helps ensure the effectiveness and availability of vision sensors.

At present, research for image dehazing in the computer vision field can be divided into three categories: (1) traditional image enhancement method, (2) fog removal based on prior image analyses, and (3) fog removal based on deep learning methods.

Earlier studies have focused on image dehazing through image enhancement techniques, such as histogram equalization [2,3], the design of specific filters [4–6], gamma correction [7,8], Retinex color theory [9–12], wavelet transform [13,14], and other methods, which highlight areas of interest in the image while suppressing or removing visual interference information. Experiments have shown that when the shooting scene has obvious depth changes or uneven fog distribution, such dehazing methods often fail because these methods simply improve the contrast of an image and lack the distinction of depth information, thereby resulting in unsatisfactory results [15].

Fog removal methods based on image priors such as dark channel prior (DCP) [16], color decay prior (CAP) [17], non-local prior [18], and color distribution line prior [19] usually estimate the restoration parameters and restore the image by using atmospheric scattering models of fog formation [20]. Some methods utilize multiple priors for image dehazing, for example, fusing contrast and exposure for image dehazing [21]. Methods based on image prior need to fully analyze the specific target scene and assumption conditions; otherwise, restored images are prone artifacts.

Dehazing methods based on deep learning can be roughly divided into two types. A series of methods are based on atmospheric scattering models and use deep networks, such as CNNs, RNNs, and GANs, to learn the restoration parameters of foggy images. For example, Ren et al. proposed a multi-scale convolutional neural network (MSCNN) [22] to estimate the transmission map. Cai et al. proposed the DehazeNet [23] network and converted the manual feature extraction method into a convolutional network layer to estimate the transmission map. By reconstructing the atmospheric scattering model, Li et al. designed an end-to-end dehazing network (AOD-Net) [24] to jointly estimate the transmission map and atmospheric light, reducing the accumulated errors in the parameter estimation process. Alona Golts et al. proposed an unsupervised transmission map prediction method for fog removal based on a context aggregation network (CAN) [25]. Yang et al. introduced a de-entanglement fog removal network (DDN) [26] to estimate scene brightness, transmittance, and global atmospheric light factor. Jiang et al. [27] estimated the depth of a scene based on an RNN algorithm to obtain the transmission map, and then restored the fog-free image through the atmospheric scattering model [28,29].

Another kind of deep-learning-based method is to learn the direct mapping from the foggy image to the clear image. For example, Zhang et al. proposed a densely connected pyramid network (DCPDN) [30], and Qu et al. proposed a pixel-to-pixel EPDN network for dehazing [31]. Methods like FD-GAN [32] and CycleGAN [33] are based on generative adversarial networks (GANs), which use a generator to generate clear images from foggy images directly; then, the authenticity of the fog removal image is judged via a discriminator. Wu et al. established an AECR-Net compact dehazing network model [34] by using a contrastive learning method; Jiang et al. proposed a deep hybrid dehazing network [35], which regards image dehazing and detail refinement as two independent tasks, and established two sub-networks for training. Dehazing methods based on vision Transformers [36] are also emerging. As this kind of method ignores the guidance information of prior features, it is prone to undesirable artifacts in the recovered results.

There are also some methods that achieve fog removal through the fusion of multiple images or multiple features of an image [37–40] or through removing fog of a single image based on additional information, such as depth information [41] and a three-dimensional model of the scene [42]. These methods require additional data information or rely on the manual selection of preprocessing methods and multiple images, and are not suitable for all scenes. According to a literature search, there is a lack of research on image dehazing

specifically for forest fog. In [43], the authors adopted methods commonly used in the field of computer vision for forest image dehazing, but the restoration results showed undesirable artifacts in the shadow areas, which are commonly seen in forest images.

By comparing the dehazing results with other methods, it can be found that if the prior information of forest images can be fully utilized, the dehazing results can be effectively improved. A dehazing algorithm based on deep learning can effectively process fog images with various fog degrees and various scenes owing to its efficient learning ability. Therefore, this study proposes combining the prior information of images with a deep learning network, thus establishing a lightweight forest scene image dehazing network that can reduce computational load while maintaining the best dehazing results.

The research methods and technological approaches of our proposed work can provide solutions and theoretical foundations for the dehazing problem in visual monitoring systems in complex agricultural and forestry environments. It can also be widely applied in outdoor visual information processing fields, including urban video surveillance, autonomous driving, rural video monitoring systems, and military purposes, indicating its broad application prospects.

The contributions of this paper are four-fold:

1. Firstly, a forest image dehazing method based on a lightweight deep learning network and joint image priors is proposed. Based on prior feature analysis, we employ the original degraded fog image and its white balance, contrast, and gamma correction feature maps as the input to the network to strengthen its ability to learn detailed image features extracted from forest scene images.
2. Secondly, our proposed dehazing network is designed based on an encoder–decoder architecture. The encoder structure consists of four different kinds of Ghost Bottleneck blocks, which adopt an SE attention mechanism, to better learn the abundant forest image features. Focusing on reducing the computational cost of the network, a lightweight upsampling module combining a bilinear interpolation method with a convolution operation is proposed and used in the decoding part, which decreases the time and computation to obtain better fog removal results.
3. Thirdly, aiming at obtaining clearer edges and better texture detail while retaining the color features of forest images, a loss function consisting of  $L_1$  loss and multi-scale structural similarity (MS-SSIM) loss is specially designed.
4. Lastly, to better solve the forest scene image dehazing problem, a forest fog image dataset is established, and our proposed network presents favorable results when compared to other methods. Our proposed method is trained on both indoor and outdoor datasets and can adapt to different dehazing scenarios.

This article is organized as follows: The background of image dehazing and the contributions of this study are presented in Section 1. The related work of the lightweight GhostNet, white balance prior, contrast prior, and gamma correction are described in Section 2. Section 3 gives a detailed description of the proposed network architecture. Section 4 verifies the effectiveness of the proposed dehazing network via abundant experiments. A conclusion is drawn in Section 5.

## 2. Related Work

### 2.1. GhostNet

In order to mitigate the computational burden in network calculations, Han, K. et al. proposed the utilization of a Ghost module as a replacement for traditional convolutional operations [44]. This Ghost module accomplishes fewer computations through two distinct steps: (1) generating feature maps with a lower number of channels using conventional convolutions; (2) further generating new feature maps through a depth-wise convolution [45], which are then merged with the feature maps produced in step (1) to obtain the final feature maps. By stacking and integrating several Ghost modules with skip connections, a lightweight network architecture, GhostNet, can be constructed.

## 2.2. Image Prior

The effectiveness of deep-learning-based image dehazing algorithms largely depends on the network's ability to extract image features. Therefore, analyzing the characteristics of foggy images in forest areas is beneficial for developing image dehazing algorithms specific to forest regions. In this section, we analyze the features of foggy images in forest areas, including white balance, contrast, and gamma correction.

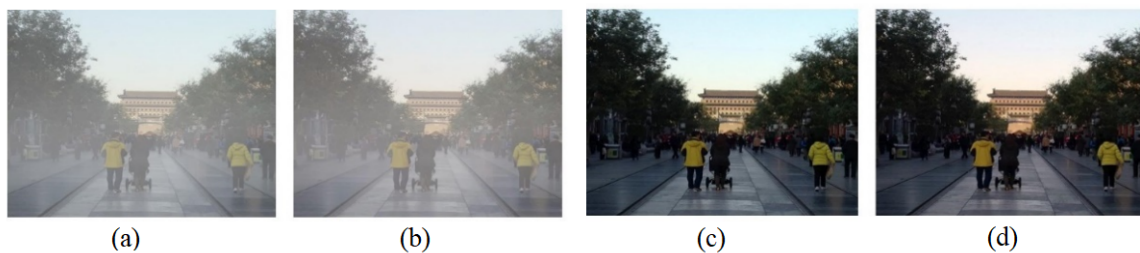
### 2.2.1. White Balance Feature

The human visual system possesses color constancy, which refers to the ability to perceive consistent colors of objects across different lighting conditions and imaging settings. However, current computer and intelligent vision systems have not yet achieved such a level of resolution. Therefore, in the case of hazy weather conditions, techniques are often employed to eliminate the color cast caused by atmospheric light in captured hazy images. One commonly used approach is to utilize white balance reference images to correct this color bias. According to the gray-world method [46], the white balance feature image is obtained through the following three steps:

1. Calculate the average values of the RGB channels— $\bar{R}$ ,  $\bar{G}$ , and  $\bar{B}$ —and then take the average again to obtain the “gray” value,  $\overline{Gray} = \frac{\bar{R} + \bar{G} + \bar{B}}{3}$ ;
2. Calculate the gain coefficients for the RGB channels,  $K_R = \frac{\overline{Gray}}{\bar{R}}$ ,  $K_G = \frac{\overline{Gray}}{\bar{G}}$ , and  $K_B = \frac{\overline{Gray}}{\bar{B}}$ ;
3. Using the von Kries diagonal model, adjust the values of the RGB components for each pixel  $y$  in the image using Equation (1), which results in the adjusted white balance feature image.

$$\begin{cases} R' = R \times K_R \\ G' = G \times K_G \\ B' = B \times K_B \end{cases} \quad (1)$$

Figure 1 shows the white balance feature map of the foggy and clear images.



**Figure 1.** Images of white balance features. (a) Fog-degraded image. (b) White balance map of (a). (c) Clear image. (d) White balance map of (c).

### 2.2.2. Contrast Features and Gamma Correction

Contrast features play a significant role in the visibility of an image. Generally, a higher contrast indicates a clearer image with richer information, while a lower contrast results in a blurred image. As shown in Figure 2, images captured in hazy weather exhibit a grayish and blurry appearance, with less distinct contours and details compared to images taken on a clear day. Tang et al. proposed the use of multi-scale local maximum contrast features to assist in the analysis of hazy images based on the dark channel prior [47]. They employed a random forest regression model to achieve the goal of haze removal.

In the GFN dehazing algorithm proposed by Ren et al. [40], the contrast enhancement feature was introduced as prior information. Along with the original hazy image, other selected features are then input into a convolutional network to obtain the weights of

various haze-related features across the entire image. Inspired by Ancuti et al. [37], Ren et al. calculated the contrast enhancement feature.  $C_{enh}$ , using Equation (2):

$$C_{enh} = \mu(I - \bar{I}) \tag{2}$$

where  $I$  represents the hazy image and  $\bar{I}$  represents its average luminance image. The coefficient  $\mu$  ( $\mu = 2 \times (0.5 + \bar{I})$ ) is utilized to linearly increase the computed luminance. As shown in Figure 2, in regions with dense fog or white objects, the values of  $(I - \bar{I})$  may dominate, leading to a phenomenon known as “darkening” in the computed contrast feature image. To address this issue, a gamma-corrected image,  $I_{gam}$ , is introduced as a supplementary contrast enhancement feature:

$$I_{gam} = \alpha I^\gamma \tag{3}$$

In Equation (3), when setting  $\alpha = 1$  and  $\gamma = 2.5$ , as shown in Figure 2, gamma correction effectively removes the black patches in the contrast enhancement feature and enhances the visibility of the image.

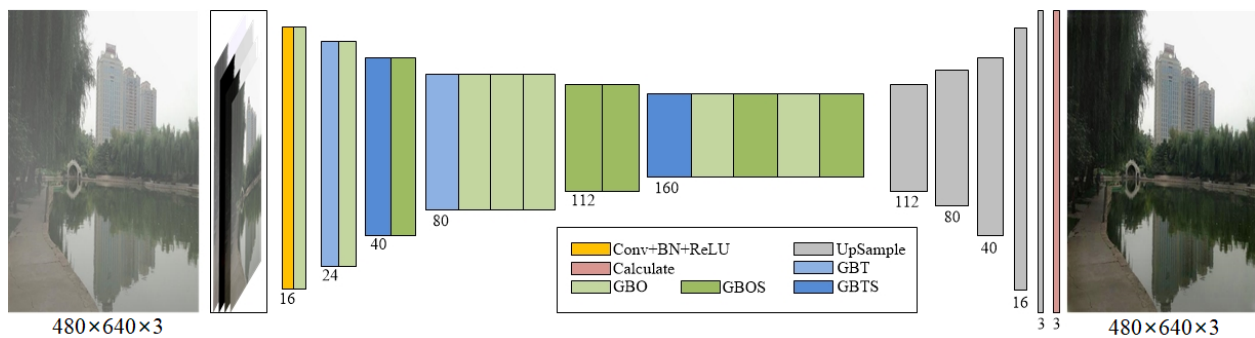


**Figure 2.** Images of contrast enhancement features and gamma correction features. (a) Image with fog. (b) Enhanced contrast features. (c) Gamma correction feature.

### 3. Proposed Lightweight Forest Scene Image Dehazing Network

#### 3.1. Network Architecture

The architecture of our proposed dehazing network is illustrated in Figure 3. For a single hazy image, we first compute its white balance feature image, contrast enhancement image, and gamma-corrected image. These prior feature images as well as the original degraded image are then fed into the proposed dehazing network. Prior information assists in updating the network weights and preserves richer detailed information. Then, the dehazing network estimates the parameters of the atmospheric scattering model, and finally, the recovered image is obtained.



**Figure 3.** The overall network architecture of the proposed method.

The proposed network utilizes convolution operations to extract image features and uses upsampling methods to restore the resolution and channel numbers of the features. As shown in Figure 3, the first layer of the network consists of a standard  $3 \times 3$  convolutional

layer followed by a batch normalization (BN) layer and an ReLU activation function. Next, four types of Ghost Bottleneck blocks are alternately stacked and used to extract image features, with differences in stride and the inclusion of a Squeeze-and-Excitation (SE) structure. This operation reduces the image resolution while increasing the number of channels. Then, five upsampling modules are designed to restore the size of the feature maps. Finally, the deformed form of the atmospheric scattering model is embedded as the calculation layer at the end of the network to compute the dehazed result.

Figure 4 illustrates the four types of Ghost Bottleneck blocks used in the network architecture, the SE structure [48], and the structure of the upsampling module. The four types of Ghost Bottleneck blocks include GBO (stride 1), GBOS (stride 1 with the SE structure), GBT (stride 2), and GBTS (stride 2 with the SE structure). GBO and GBT structures are similar, and both consist of a Ghost Module sub-block and a shortcut connection. GBT also includes an additional depth-wise convolution layer, which is an efficient operation in lightweight networks for handling spatial information. The shortcut connection operation merges some features, allowing the network to preserve more detailed information. The Ghost Module is composed of the concatenation of the outputs of two convolution operations, extracting features with fewer computations. GBOS and GBTS are based on GBO and GBT, respectively, with the addition of the SE structure. The SE structure is an attention mechanism that assigns different weights to each feature, enabling the network to focus on more important features. The implementation is as follows: Firstly, the feature map with a size of  $H \times W \times C$  is globally pooled to size  $1 \times 1 \times C$  and then passed through a fully connected layer to predict the importance of each channel, obtaining the weights of each channel. This is followed by another fully connected layer to adjust the feature map and multiply it with the input feature, resulting in the final output. The SE structure assigns different weights to each feature map, focusing on more important features, as shown in Figure 4f. The upsampling module consists of a bilinear interpolation upsampling layer and two convolution layers. The former restores the resolution of the image, while the latter restores the number of channels.

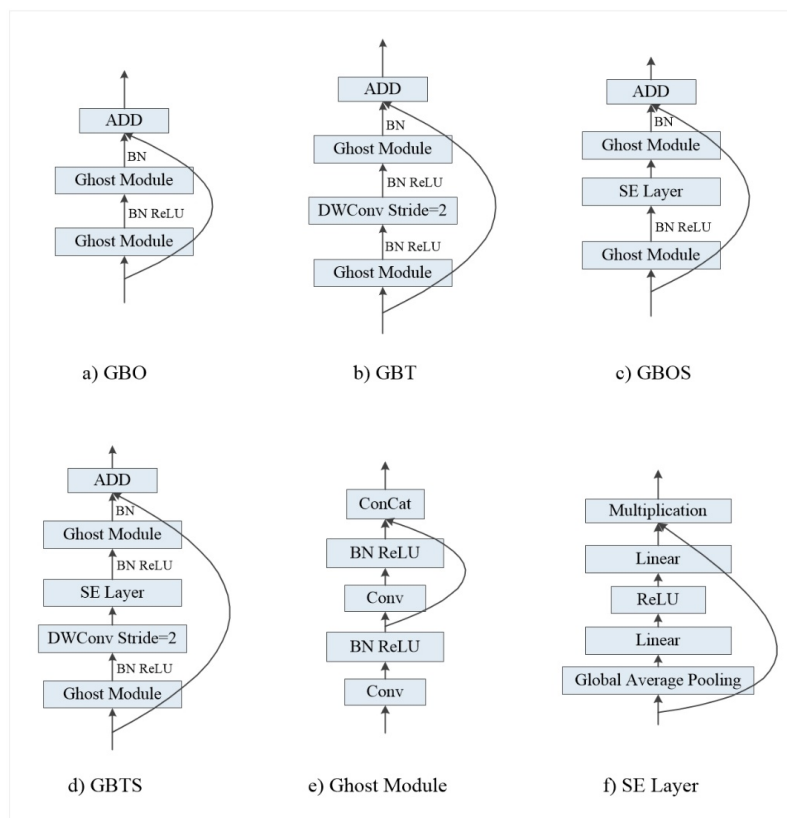


Figure 4. The structure of different modules. The big arrow represents the shortcut connection.

### 3.2. Loss Function

#### 3.2.1. $L_1$ Loss Function

$L_1$  loss function, also known as least absolute error, is defined as shown in Equation (4);  $y(i)$  and  $y^*(i)$  represent the restored image and the ground truth image, respectively; and  $i$  represents each pixel of the image.

$$L_1 = \frac{1}{n} \sum_{i=1}^n |y(i) - y^*(i)| \quad (4)$$

The  $L_1$  loss function better reflects the error between the ground truth and predicted values. A smaller loss value indicates the higher precision of the predicted values using the model. The derivative of this error, when used with image inputs during back propagation, is calculated at each pixel. This usually does not lead to gradient explosion issues and has higher robustness. For image restoration models, the  $L_1$  loss function is more accepting of outliers and is less smooth. This allows it to capture more texture features of the restored image, resulting in a more natural visual effect.

#### 3.2.2. MS-SSIM Loss Function

Structural Similarity (SSIM) is a metric used to measure the similarity between two images. It incorporates information about the brightness, contrast, and structural details of the images, making it more aligned with human visual perception. Generally, models trained with the SSIM loss function produce restoration results that have more details compared to  $L_1$  or  $L_2$  losses, but they may introduce noise in the boundary areas. However, the Multi-Scale Structural Similarity (MS-SSIM) loss function can address most of the boundary noise issues. It takes into account the influence of image resolution. It repeatedly downsamples the restored image and the ground truth image, generating multiple images with different resolutions (usually five images). Then, it evaluates the SSIM index for the images with different resolutions and combines the evaluation values, as shown in Equation (5). Zhao et al. [49] and Huang et al. [50] have explained and utilized this loss function to improve the boundary information of images and other aspects.

$$L_{MS-SSIM} = 1 - \prod_{m=1}^M \left( \frac{2\mu_p\mu_g + C_1}{\mu_p^2 + \mu_g^2 + C_1} \right)^{\beta_m} \left( \frac{2\sigma_{pg} + C_2}{\sigma_p^2 + \sigma_g^2 + C_2} \right)^{\gamma_m} \quad (5)$$

In Equation (5),  $\mu_p$ ,  $\mu_g$ ,  $\sigma_p$ , and  $\sigma_g$  represent the mean and variance of the restored image and the ground truth image, while  $\sigma_{pg}$  represents the covariance between the restored image and the ground truth image.  $\beta_m$  and  $\gamma_m$  indicate the importance of mean and variance in each scale, respectively.  $C_1$  and  $C_2$  are small constants used to avoid division by zero. In practical applications, in order to achieve higher computational efficiency, Gaussian kernel convolution is commonly used to calculate the mean, variance, and covariance of images.

#### 3.2.3. Total Loss Function

SSIM-based loss functions are not sensitive to uniform biases, which can result in changes in color and brightness, making the restored results darker. However, they can preserve high-frequency information. However,  $L_1$  loss has a better color preservation ability. Therefore, in order to reduce the deviation in image brightness and color while preserving more and better edge and detail information, we adopted a combination of  $L_1$  loss and MS-SSIM loss to train the network. Zhao et al. provided the calculation formula for this combined loss function [49]:

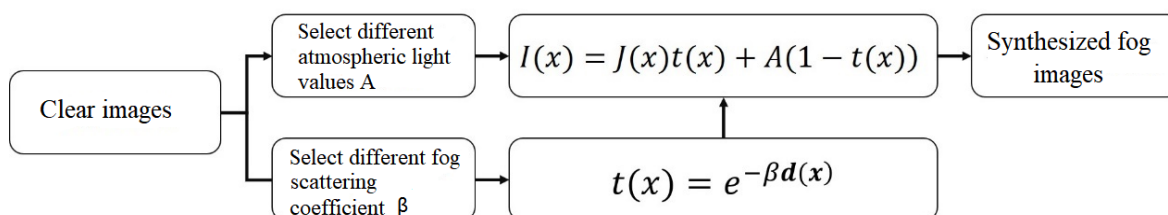
$$L_A = \alpha \cdot L_{MS-SSIM} + (1 - \alpha) \cdot G_{\sigma_G^M} \cdot L_1 \quad (6)$$

In Equation (6),  $G_{\sigma_G^M}$  represents Gaussian filters with standard deviation,  $\sigma_G$ , used at  $M$  different scales.  $\alpha$  represents the weighting factor of the two losses in the total loss

( $\alpha = 0.025$ ). In the context of image dehazing, the emphasis is on the overall restoration of the reconstructed image. It is important to ensure the clarity of object boundaries, textures, and other details while maintaining the natural color tone of the image as a whole. In theory, the joint loss of  $L_1$  and MS-SSIM can effectively address the dehazing challenge mentioned above. Through experimental verification, this combined loss has shown promising results in improving the visibility of hazy images.

### 3.3. Dataset and Training Parameters

To train the proposed network, similar to the existing deep-learning-based fog removal methods, synthetic datasets are used for training. Synthesized datasets are generated based on atmospheric scattering models [20] using clear images with corresponding depth maps. A flowchart of the synthetic dataset is shown in Figure 5.



**Figure 5.** A flowchart of synthetic dataset obtained.

In Figure 5,  $I$  is the generated hazy image,  $J$  is the original clear image,  $d$  is the depth map of an image, and  $t$  is the transmission map calculated via the depth map,  $d$ , and scattering coefficients,  $\beta$ .

In order to apply dehazing to different scenarios, our proposed dehazing network is trained both on indoor and outdoor hazy datasets. For indoor scenes, the NYU2 Depth Database [51] was used to synthesize hazy images. The NYU2 depth dataset contains 1449 images from 464 different indoor scenes. We randomly selected 109 images from different scenes to generate the testing set and the remaining 1340 images were used to generate the training set. By setting atmospheric light values  $A \in \{0.6, 0.7, 0.8, 0.9, 1.0\}$  and scattering coefficients as  $\beta \in \{0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6\}$ , the synthesized indoor hazy image set containing 46,900 hazy–clear image pairs was obtained for training via the method shown in Figure 5. For the testing set, we randomly selected two values of  $A$  and three values of  $\beta$  and, thus, generated 654 images for testing. For outdoor scenes, we directly used the RESIDE OTS synthetic dataset [52] for training, which is a sub-dataset of the RESIDE database containing 72,135 synthetic hazy images. For testing, the RESIDE SOTS dataset, which consists of 500 hazy images, was used.

We implement our proposed lightweight forest scene dehazing network in distribution mode with two 1080ti GPUs under the Pytorch framework. The network parameters are initialized using the Kaiming method proposed by He et al. [53]. The ADAM optimization algorithm is used to train the network, which has four parameters to set, including learning rate, exponential decay rate of the first-order moment estimate ( $\beta_1$ ), exponential decay rate of the second-order moment estimate ( $\beta_2$ ), and a parameter  $\epsilon$  to avoid the occurrence of 0. The training parameters are provided in Table 1. We train the models for 30 epochs separately on the indoor and outdoor training sets. The models begin to converge around the 5th epoch, and the best prediction results are obtained at the 11th and 10th epochs for the indoor and outdoor datasets, respectively.



**Table 1.** Training parameter setting of the proposed method.

Parameter	Value
Learning rate	0.0001
Exponential decay rate of the first-order moment estimate ( $\beta_1$ )	0.9
Exponential decay rate of the second-order moment estimate ( $\beta_2$ )	0.999
Parameter $\epsilon$	$1 \times 10^{-8}$
Batch size	4
Training steps	30
Training image resolution	$480 \times 640$

#### 4. Results

In order to evaluate the performance of the proposed network, four types of test sets are used for the experiments: (1) NYU2 indoor synthetic hazy images; (2) RESIDE OTS outdoor synthesized hazy images; (3) RTTS real-world hazy images; and (4) real-world forestry scene hazy images. The first two datasets of fog images have their corresponding clear images as the ground truth, so we quantitatively analyze the fog removal results by using objective evaluation indicators, and visibility quality analyses of the proposed method compared to other state-of-the-art methods are also given. The latter two datasets of fog images are real-world fog images, and there is no corresponding clear image, so we compare the results visually.

Peak Signal to Noise Ratio (PSNR) and Structural Similarity (SSIM) are utilized to evaluate the quality of the dehazed images in our experiments, which are the most widely used objective indices. PSNR is an evaluation index based on the difference between the corresponding pixels of two images. The larger the value, the better the dehazing results. SSIM is used to fit the real perception of human beings and evaluate image quality from three aspects, including brightness, contrast, and structure. The value range is from 0 to 1. The larger the value, the better the dehazing result, and a value of 1 indicates that the two images evaluated are identical. Moreover, single image processing time is compared to evaluate the speed of different methods.

##### 4.1. NYU2 Test Set

Figure 6 shows the dehazing results of the proposed model and seven other state-of-the-art dehazing algorithms on the synthesized indoor test set of NYU2. Based on subjective visual evaluation, the DCP algorithm does not fully remove the haze because it was primarily designed for outdoor hazy images. It introduces a parameter to preserve some haze in distant objects to maintain depth perception. The AMEF algorithm enhances the overall color of the scene, resulting in rich and clear details in the dehazed results. However, it tends to darken certain areas of the ground and walls, which appear unnatural, such as the white floor in the first image and the walls in the fourth image. The DehazeNet algorithm improves image contrast but loses some fine texture details. The enhanced EPDN algorithm produces results that are close to the ground truth images with natural colors and clear textures. The AOD algorithm also produces darker colors and less prominent texture information but has a fast processing speed. The GFN algorithm exhibits large black patches in some white areas, such as the white wall region in the fourth image. Both the GCA algorithm and our proposed method provide reasonable color and texture information. However, the GCA algorithm falls short in handling the brown tabletop and floor area, resulting in a visual sense of being partially obscured by a thin haze, as shown in the third image. In comparison, the results of the proposed algorithm visually resemble the ground truth images and exhibit a more natural restoration of details and color information. Additionally, as shown in Table 2, the proposed algorithm performs best in terms of PSNR and SSIM indices, and the restoration time for a single hazy image is within an acceptable range.

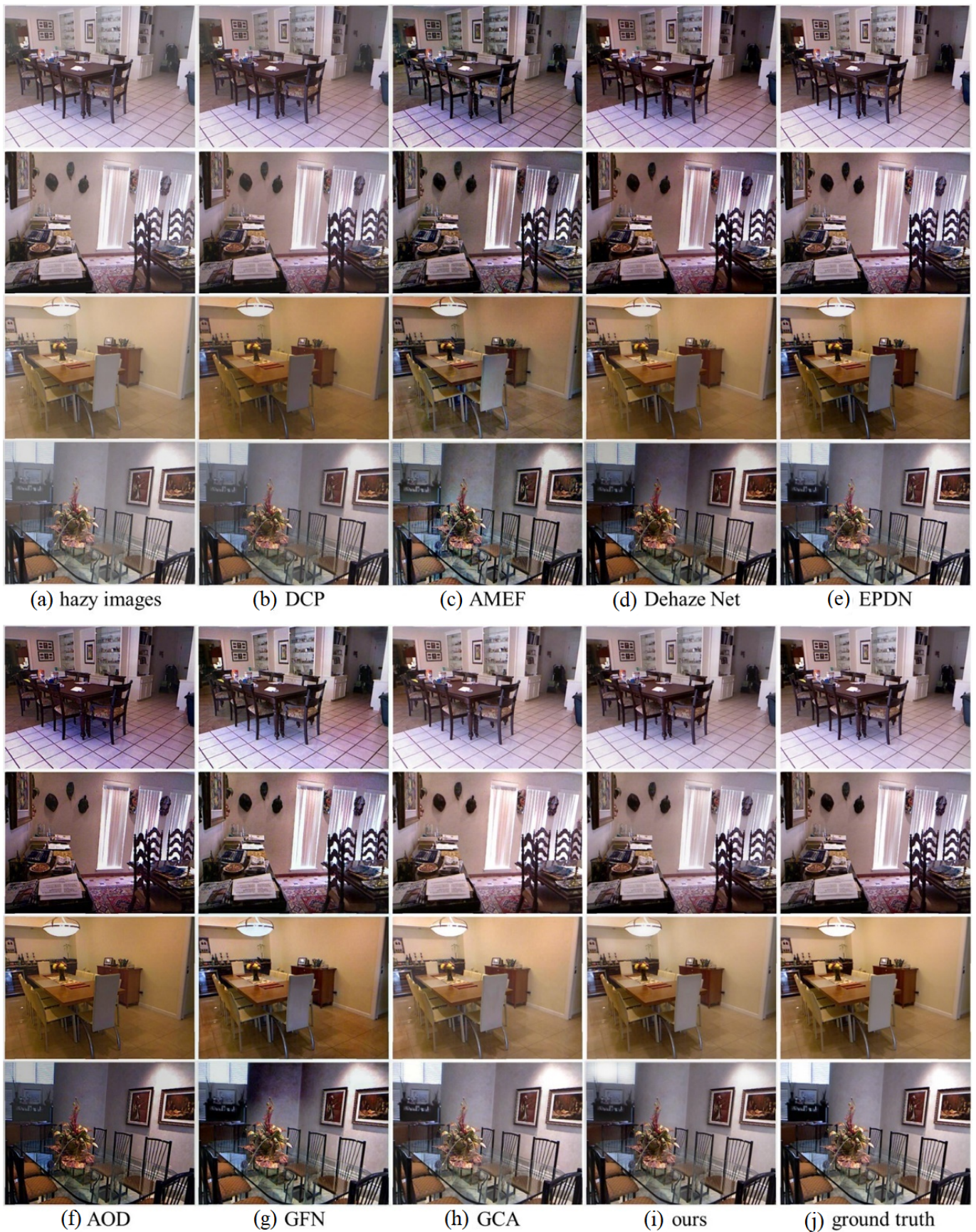


Figure 6. Dehazing results on NYU2 testing set.

**Table 2.** The quantitative comparison on NYU2 testing set. The best results are in bold.

Method/Evaluation Criteria	PSNR	SSIM	Single-Image Processing Time (s)
DCP [16]	21.7767	0.9234	0.44
AMEF [54]	17.9960	0.8839	0.74
Dehaze Net [23]	20.5656	0.9405	2.00
EPDN [31]	25.2795	0.9513	0.27
AOD [24]	18.9208	0.9190	0.04
GFN [40]	23.0912	0.9469	5.15
GCA [55]	25.2486	0.9579	0.19
<b>Ours</b>	<b>26.0063</b>	<b>0.9594</b>	0.26

#### 4.2. RESIDE OTS Test Set

The algorithms shown in Figure 7 all perform haze removal on synthesized outdoor hazy images. Among them, the results of the proposed algorithm are visually the most similar to the ground truth images, with reasonable and clear restoration of texture, edges, and color details of trees and buildings. The DCP algorithm preserves haze for distant objects to maintain depth perception, and it also introduces little distortion in the majority of sky areas in this test set. The AMEF algorithm focuses on enhancing the edge details of objects. By observing the second and third images, it can be seen that the algorithm deepens the edges and significantly improves image contrast. The classical dehazing network DehazeNet and the enhanced EPDN network both produce darker results and lose a lot of texture details, especially in the dark areas of the original hazy images, such as the lower area of the pedestrian bridge in the first image and the trees in the second and fourth images, which are processed into completely black patches. The results of the AOD algorithm are similar to the algorithm proposed in this section, but the colors in the AOD algorithm are more vibrant. The GFN algorithm produces black regions due to its overly dark handling of object colors, blurring the original texture details, and its color restoration for distant buildings is unreasonable, such as the walls of tall buildings in the fourth image. The GCA algorithm usually achieves good dehazing results, but there presents unrealistic colors in sky areas. To quantify the evaluation, the PSNR and SSIM values are compared, as well as the processing time for a single hazy image, using different dehazing methods. As shown in Table 3, our proposed method achieves the best metrics, and its processing speed for a single hazy image is also competitive.

**Table 3.** The quantitative comparison on OTS testing set. The best results are in bold.

Method/Evaluation Criteria	PSNR	SSIM	Single-Image Processing Time (s)
DCP [16]	21.3657	0.8968	0.45
AMEF [54]	17.7227	0.8683	0.78
Dehaze Net [23]	22.2356	0.9219	2.05
EPDN [31]	22.5700	0.8630	0.27
AOD [24]	19.5119	0.9104	0.06
GFN [40]	22.3000	0.8800	5.25
GCA [55]	22.2013	0.9280	0.18
<b>Ours</b>	<b>25.5830</b>	<b>0.9360</b>	0.26



Figure 7. Dehazing results on OTS testing set.

#### 4.3. RTTS Test Set

RTTS image set is a real-world hazy image set containing 4322 images without ground truth images, which is used as a test set for object detection under bad weather in the image processing field. As shown in Figure 8, we compare and analyze the dehazing results of various dehazing algorithms on real outdoor foggy images on the RTTS image sets. The dehazing results of the DCP algorithm present obvious fog residuals, and distant objects are more difficult to restore. The AMEF algorithm emphasizes edge information in the image, especially for nearby edges, but there are also obvious areas with fog residuals, such as the tree area in the second and third images. The results of DehazeNet and EPDN algorithms are darker and lose some details, such as DehazeNet's processing of the person in the first image and EPDN's treatment of tree crowns in the third image. The AOD algorithm shows a loss of details and incomplete dehazing in some areas. This may be due to complex scenes with thick haze in certain regions, where the lightweight AOD convolutional network does not extract detailed image features sufficiently. The GFN algorithm constructs a network based on prior information about white balance and contrast from hazy images. However, according to our analysis of contrast features, contrast feature images tend to exhibit black patches more prominently. If the relationship between hazy features and haze-free images is improperly learned via the network, many black pixels will appear in its restoration results, leading to a loss of details. The dehazing results obtained via the GCA algorithm are similar to those achieved via our proposed method, with clearer dehazing results, and the overall color appears more natural. The UHD algorithm produces brighter images with a greater restoration of details and textures, but occasionally, it has difficulties in removing distant fog, such as the middle region depicted in the second figure. It is worth mentioning that there exist certain differences between real hazy images and synthesized ones, and deep-learning-based algorithms rely more on the types of scenes and fog conditions in the training dataset. By comparison, our proposed dehazing algorithm recovers most of the detailed texture and color information, and the results are more natural visually.

#### 4.4. Forestry Test Dataset

We collected 212 real forest scene foggy images to test the performance of the proposed algorithm. Figure 9a presents the real-world foggy image with light mist, while Figure 9b–j show the results of nine commonly used defogging algorithms, including our proposed method. According to the dehazing results, the DCP algorithm is more natural and closer to the color of the original image. The results of the AMEF algorithm are colorful, show rich details, and emphasize the edge information. The DehazeNet results are close to the DCP algorithm, but the trunk part is darker. The EPDN algorithm treats tree trunks darkly, resulting in the loss of some detail, but it shows better performance for distant fog, such as the woods afar in the second and third images. Some details are lost in the results of both AOD and GFN algorithms, such as the trunk area in the first image and the tree crown processing in the fourth image when using the AOD algorithm. The results of the GCA algorithm and our proposed algorithm are more similar in color. The former is better for the texture recovery of the trunk, and the latter is slightly darker. The UHD algorithm obtains the brightest picture, but because of the brightness, some areas still have the visual effect of being covered by mist, such as the first and second images.

The experimental results show that the proposed algorithm is effective for both indoor and outdoor composite fog images, real-world fog images, and real forestry fog images, and can obtain restored images with rich colors, clear textures, and more natural visual effects.



Figure 8. Dehazing results on RTTS testing set.

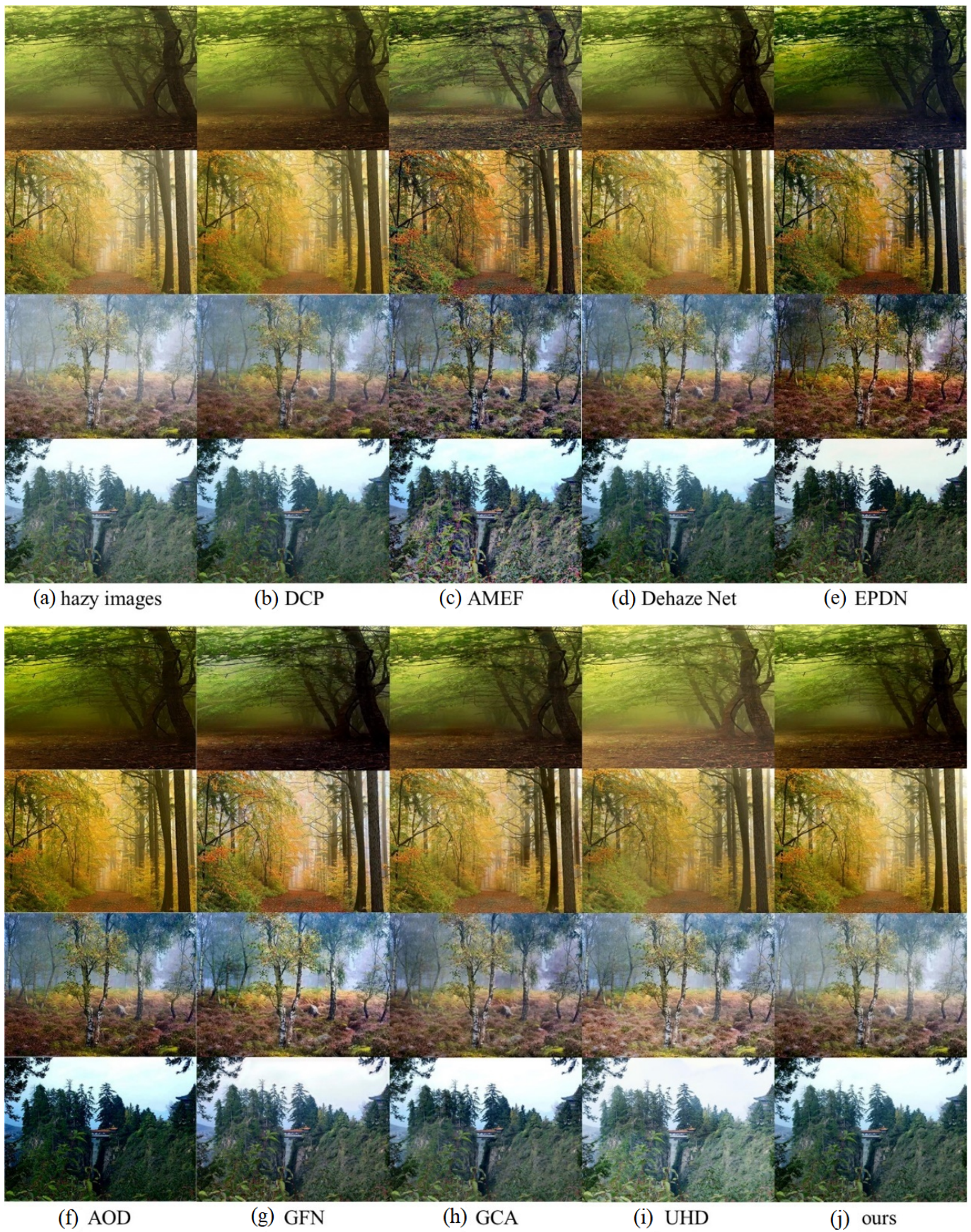


Figure 9. Dehazing results on forestry testing dataset.

#### 4.5. The Impact of Prior Information on Model Performance

The main purpose of image dehazing is to make the details, edges, textures, colors, and other features of the image clearer without making significant changes to the overall tone, making the image appear more natural. Therefore, feature extraction from the image is particularly important. The better and more detailed the feature information extracted and preserved by the model, the clearer and more visually appealing the dehazed result will be. Considering the network's extraction of relevant features, i.e., the updating of the weight parameters in the network, the following controversy arises: which approach, learning features directly from hazy images or learning features from hazy images combined with prior information (known features), can yield better dehazing results? This section verifies the model's requirements for input information through experiments. Section 2.2 of this paper describes several commonly used haze-related features, which can be used as prior information to assist network learning. The GFN image dehazing network adopts feature fusion, indirectly proving the importance of contrast features and white balance features. In order to reduce computational complexity for the entire dehazing algorithm, this section chooses white balance, contrast, and gamma correction prior features, along with hazy images as inputs, then compares and analyzes the results with single-input hazy images. In the experiment, the dataset, optimization algorithm, loss function, etc., are unchanged.

Table 4 presents the comparison results between single-input and multi-input approaches. The results show that using prior information to assist the dehazing network in feature extraction has certain improvement effects.

**Table 4.** The effectiveness of prior features. The best results are in bold.

Prior Information/Evaluation Index	PSNR	SSIM	Single-Image Processing Time (s)
w/o	22.5643	0.9210	0.23
<b>w</b>	<b>26.0063</b>	<b>0.9594</b>	<b>0.26</b>

#### 4.6. The Impact of Upsampling Methods on Model Performance

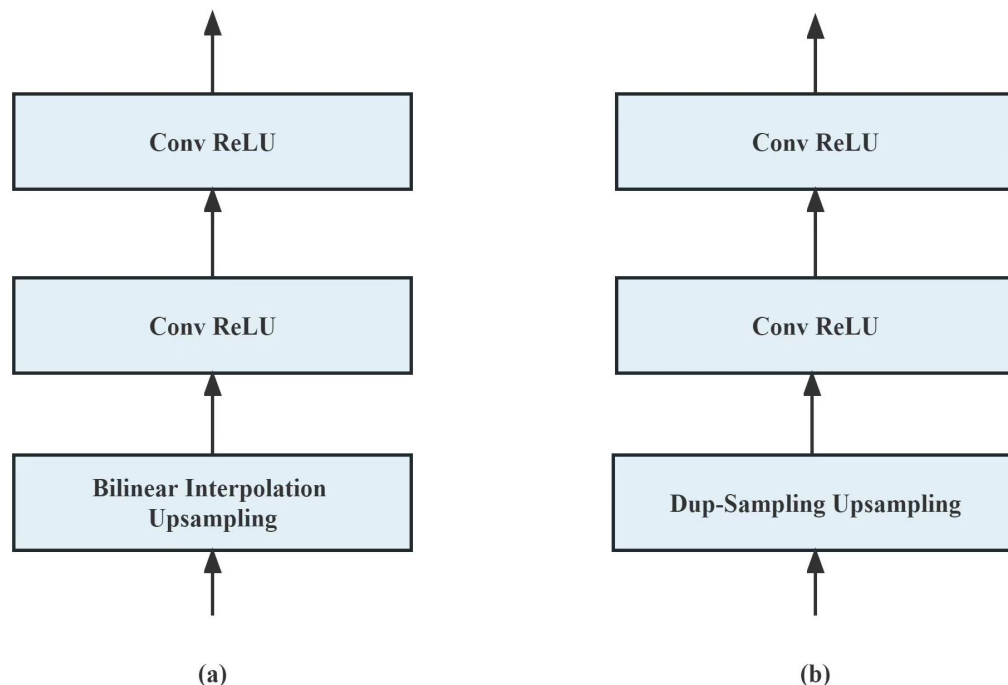
This section aims to discuss the impact of different upsampling methods on the performance of the proposed model. Methods such as deconvolution, interpolation (upsampling) + convolution, and Dup-sampling [56] + convolution can all restore the feature map to the same size as the input image. However, deconvolution, being the inverse process of convolution, requires careful parameter configuration (stride, etc.), as improper configuration can lead to the appearance of a checkerboard grid artifact. Therefore, this section only focuses on an experimental comparison of the interpolation + convolution and Dup-sampling + convolution upsampling methods.

The interpolation + convolution upsampling method uses interpolation, such as bilinear interpolation or nearest-neighbor interpolation, to restore the image resolution and uses convolution to restore the channel number of the image. In this section, bilinear interpolation + two convolutions are used to construct the upsampling module, as shown in Figure 10a.

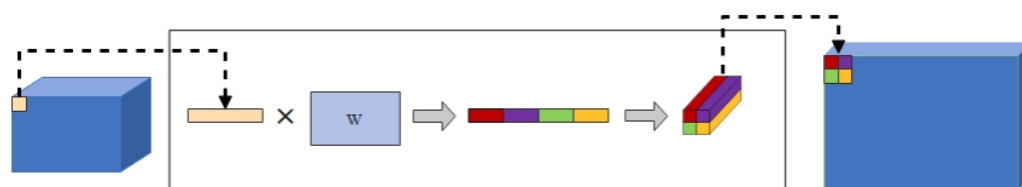
The Dup-sampling upsampling method proposed by Tian et al. can simultaneously restore the resolution and channel number of the feature map [56]. As shown in Figure 11, considering an input feature matrix of size  $H \times W \times C$ , this input matrix is first divided into  $H \times W$  small matrices of size  $1 \times C$ . Then, each of these  $1 \times C$  small matrices is multiplied by a matrix  $W$  of size  $C \times N$ , resulting in several  $1 \times N$  matrices (where  $W$  is obtained through training). Next, these  $1 \times N$  matrices are rearranged and combined to form several  $2 \times 2 \times N/4$  matrices, where each  $2 \times 2 \times N/4$  small matrix represents a horizontal stretch and vertical compression operation. Finally, these several  $2 \times 2 \times N/4$  small matrices are neatly arranged together, completing one round of Dup-sampling. As shown in Figure 10b, in this section, Dup-sampling is used together with two convolutions to form an upsampling module. Experimental analysis and comparison are conducted between



this method and the interpolation + convolution method while keeping the training set, optimization method, and test set unchanged. Table 5 presents the comparison results of the two upsampling methods. The results indicate that the interpolation upsampling method using bilinear interpolation produces better haze removal results.



**Figure 10.** The structure of different upsampling modules. (a) Conv + upsampling structure (b) Dup-sampling structure.



**Figure 11.** The structure of Dup-sampling method.

**Table 5.** The results of different upsampling methods. The best results are in bold.

Upsampling Information/ Evaluation Index	PSNR	SSIM	Single-Image Processing Time (s)
Dup-sampling	23.6496	0.9426	0.28
<b>Conv + upsample</b>	<b>26.0063</b>	<b>0.9594</b>	<b>0.26</b>

## 5. Conclusions

In this study, we propose a lightweight forest scene image dehazing network based on joint image priors. Based on the prior feature analysis of fog-free and foggy images of forest scenes, the degraded foggy image and its three important feature maps, including white balance, contrast, and gamma correction, are used as inputs to the network to assist the network in learning and extracting more detailed image features. To better facilitate the deployment of dehazing networks in forestry edge devices, a lightweight deep network is designed. Four different kinds of Ghost Bottleneck blocks, which adopt an SE attention mechanism to better learn the abundant forest image features for our network, are designed. Focusing on reducing the computational cost of the network, a lightweight upsampling module combining a bilinear interpolation method and a convolution operation is proposed

and used in the decoding part, which requires less time and computation to obtain better fog removal results. In order to better retain the texture features of the restored images and make the recovery image visually more natural, a loss function combining  $L_1$  loss function and multi-scale structural similarity (MS-SSIM) is adopted. As a result, the network can learn more abundant image features and, thus, improve the overall image dehazing capability in forest scene images.

The experimental results on both indoor and outdoor datasets demonstrated that the proposed method achieves better performance when compared to other state-of-the-art dehazing methods for forest scene images. However, like most of the existing methods, there will be a certain degree of residual haze in thick hazy regions for real-world foggy images. In our future work, we will further exploit the relevant prior features of hazy forest scene images, improve our algorithm, and explore deep learning networks to better solve the problem of residual haze in thick hazy regions of an image.

**Author Contributions:** Conceptualization, J.K.; methodology, X.Z., Y.M. and J.K.; software, X.Z., Y.M. and J.Z.; validation, Y.M. and X.Z.; formal analysis, Y.M. and J.K.; investigation, Y.M.; resources, X.Z. and J.K.; data curation, Y.M.; writing—original draft preparation, X.Z., Y.M., Z.J. and J.Z.; writing—review and editing, X.Z. and J.K.; visualization, Y.M. and Z.J.; supervision, J.K.; project administration, J.K.; funding acquisition, X.Z. and J.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Guangdong Basic and Applied Basic Research Foundation (grant no.: 2022A1515110024) and the Beijing Natural Science Foundation (grant no.: 6232029).

**Data Availability Statement:** The data obtained in this study will be shared upon reasonable request to the corresponding author.

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

## References

1. Shakeri, M.; Dezfoulian, M.H.; Khotanlou, H.; Barati, A.; Masoumi, Y. Image contrast enhancement using fuzzy clustering with adaptive cluster parameter and sub-histogram equalization. *Digit. Signal Process.* **2017**, *62*, 224–237.
2. Stark, J.A. Adaptive image contrast enhancement using generalizations of histogram equalization. *IEEE Trans. Image Process.* **2000**, *9*, 889–896. [[CrossRef](#)] [[PubMed](#)]
3. Xiao, L.; Li, C.; Wu, Z.; Wang, T. An enhancement method for X-ray image via fuzzy noise removal and homomorphic filtering. *Neurocomputing* **2016**, *195*, 56–64. [[CrossRef](#)]
4. Kaplan, N.H.; Erer, I. Bilateral filtering-based enhanced pansharpener of multispectral satellite images. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 1941–1945. [[CrossRef](#)]
5. He, K.; Sun, J.; Tang, X. Guided image filtering. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 1397–1409. [[CrossRef](#)]
6. Zheng, Z.; Ren, W.; Cao, X.; Hu, X.; Wang, T.; Song, F.; Jia, X. Ultra-high-definition image dehazing via multi-guided bilateral learning. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; IEEE: New York, NY, USA, 2021; pp. 16180–16189.
7. Ju, M.; Ding, C.; Guo, Y.J.; Zhang, D. IDGCP: Image dehazing based on gamma correction prior. *IEEE Trans. Image Process.* **2019**, *29*, 3104–3118. [[CrossRef](#)]
8. Ju, M.; Ding, C.; Zhang, D.; Guo, Y.J. Gamma-correction-based visibility restoration for single hazy images. *IEEE Signal Process. Lett.* **2018**, *25*, 1084–1088. [[CrossRef](#)]
9. Wang, Y.; Wang, H.; Yin, C.; Dai, M. Biologically inspired image enhancement based on Retinex. *Neurocomputing* **2016**, *177*, 373–384. [[CrossRef](#)]
10. Chen, H.; Chen, R.; Ma, L.; Li, N. Single-image dehazing via depth-guided deep retinex decomposition. *Vis. Comput.* **2022**, 1–13. [[CrossRef](#)]
11. Fu, Q.; Jung, C.; Xu, K. Retinex-based perceptual contrast enhancement in images using luminance adaptation. *IEEE Access* **2018**, *6*, 61277–61286. [[CrossRef](#)]
12. Li, P.; Tian, J.; Tang, Y.; Wang, G.; Wu, C. Deep retinex network for single image dehazing. *IEEE Trans. Image Process.* **2020**, *30*, 1100–1115. [[CrossRef](#)] [[PubMed](#)]
13. Farge, M. Wavelet transforms and their applications to turbulence. *Annu. Rev. Fluid Mech.* **1992**, *24*, 395–458. [[CrossRef](#)]
14. Rong, Z.; Jun, W.L. Improved wavelet transform algorithm for single image dehazing. *Optik* **2014**, *125*, 3064–3066. [[CrossRef](#)]
15. Wu, D.; Zhu, Q. The Latest Research Progress of Image Dehazing. *Acta Autom. Sin.* **2015**, *41*, 221–239.
16. He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *33*, 2341–2353. [[PubMed](#)]

17. 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.
18. Berman, D.; Avidan, S. 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.
19. Fattal, R. Dehazing using color-lines. *ACM Trans. Graph. (TOG)* **2014**, *34*, 1–14. [[CrossRef](#)]
20. Narasimhan, S.G.; Nayar, S.K. Interactive (de) weathering of an image using physical models. In Proceedings of the IEEE Workshop on Color and Photometric Methods in Computer VISION, Nice, France, 13–16 October 2003; Volume 6, p. 1.
21. Liu, X.; Li, H.; Zhu, C. Joint contrast enhancement and exposure fusion for real-world image dehazing. *IEEE Trans. Multimed.* **2021**, *24*, 3934–3946.
22. Ren, W.; Liu, S.; Zhang, H.; Pan, J.; Cao, X.; Yang, M.H. Single image dehazing via multi-scale convolutional neural networks. In Proceedings of the Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Proceedings, Part II 14; Springer: Berlin/Heidelberg, Germany, 2016; pp. 154–169.
23. Cai, B.; Xu, X.; Jia, K.; Qing, C.; Tao, D. Dehazenet: An end-to-end system for single image haze removal. *IEEE Trans. Image Process.* **2016**, *25*, 5187–5198. [[CrossRef](#)]
24. 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.
25. Golts, A.; Freedman, D.; Elad, M. Unsupervised single image dehazing using dark channel prior loss. *IEEE Trans. Image Process.* **2019**, *29*, 2692–2701. [[CrossRef](#)]
26. Yang, X.; Xu, Z.; Luo, J. Towards perceptual image dehazing by physics-based disentanglement and adversarial training. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32.
27. Jiang, X.; Sun, J.; Li, C.; Ding, H. Video image defogging recognition based on recurrent neural network. *IEEE Trans. Ind. Inform.* **2018**, *14*, 3281–3288. [[CrossRef](#)]
28. Zheng, F.; Wang, X.; He, D.; Li, N.; Fu, Y.; Yuan, S. Survey of Single Image Defogging Algorithm. *J. Comput. Eng. Appl.* **2022**, *58*.
29. Huang, L.Y.; Yin, J.L.; Chen, B.H.; Ye, S.Z. Towards unsupervised single image dehazing with deep learning. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019; IEEE: New York, NY, USA, 2019; pp. 2741–2745.
30. 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, 18–22 June 2018; pp. 3194–3203.
31. 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.
32. Dong, Y.; Liu, Y.; Zhang, H.; Chen, S.; Qiao, Y. FD-GAN: Generative adversarial networks with fusion-discriminator for single image dehazing. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 9–11 February 2020; Volume 34, pp. 10729–10736.
33. Engin, D.; Genç, A.; Kemal Ekenel, H. Cycle-dehaze: Enhanced cyclegan for single image dehazing. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 825–833.
34. Wu, H.; Qu, Y.; Lin, S.; Zhou, J.; Qiao, R.; Zhang, Z.; Xie, Y.; Ma, L. Contrastive learning for compact single image dehazing. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 10551–10560.
35. Jiang, N.; Hu, K.; Zhang, T.; Chen, W.; Xu, Y.; Zhao, T. Deep hybrid model for single image dehazing and detail refinement. *Pattern Recognit.* **2023**, *136*, 109227. [[CrossRef](#)]
36. Song, Y.; He, Z.; Qian, H.; Du, X. Vision Transformers for Single Image Dehazing. *IEEE Trans. Image Process.* **2023**, *32*, 1927–1941. [[CrossRef](#)]
37. Ancuti, C.O.; Ancuti, C. Single image dehazing by multi-scale fusion. *IEEE Trans. Image Process.* **2013**, *22*, 3271–3282. [[CrossRef](#)]
38. Zhao, D.; Xu, L.; Yan, Y.; Chen, J.; Duan, L.Y. Multi-scale optimal fusion model for single image dehazing. *Signal Process. Image Commun.* **2019**, *74*, 253–265. [[CrossRef](#)]
39. Guo, X.; Yang, Y.; Wang, C.; Ma, J. Image dehazing via enhancement, restoration, and fusion: A survey. *Inf. Fusion* **2022**, *86*, 146–170. [[CrossRef](#)]
40. Ren, W.; Ma, L.; Zhang, J.; Pan, J.; Cao, X.; Liu, W.; Yang, M.H. Gated fusion network for single image dehazing. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 3253–3261.
41. Hautière, N.; Tarel, J.P.; Aubert, D. Towards fog-free in-vehicle vision systems through contrast restoration. In Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 18–23 June 2007; IEEE: New York, NY, USA, 2007; pp. 1–8.
42. Kopf, J.; Neubert, B.; Chen, B.; Cohen, M.; Cohen-Or, D.; Deussen, O.; Uyttendaele, M.; Lischinski, D. Deep photo: Model-based photograph enhancement and viewing. *ACM Trans. Graph. (TOG)* **2008**, *27*, 1–10. [[CrossRef](#)]
43. Lu, J. Study on Fog Removal Algorithm for Single Forest Images. Master’s Thesis, Northeast Forestry University, Harbin, China, 2018.
44. Han, K.; Wang, Y.; Tian, Q.; Guo, J.; Xu, C.; Xu, C. Ghostnet: More features from cheap operations. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 1580–1589.

45. Howard, A.; Sandler, M.; Chu, G.; Chen, L.C.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V.; et al. Searching for mobilenetv3. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Long Beach, CA, USA, 15–20 June 2019; pp. 1314–1324.
46. Reinhard, E.; Adhikhmin, M.; Gooch, B.; Shirley, P. Color transfer between images. *IEEE Comput. Graph. Appl.* **2001**, *21*, 34–41. [[CrossRef](#)]
47. Tang, K.; Yang, J.; Wang, J. Investigating haze-relevant features in a learning framework for image dehazing. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 2995–3000.
48. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 7132–7141.
49. Zhao, X.; Wang, K.; Li, Y.; Li, J. Deep fully convolutional regression networks for single image haze removal. In Proceedings of the 2017 IEEE Visual Communications and Image Processing (VCIP), St. Petersburg, FL, USA, 10–13 December 2017; IEEE: New York, NY, USA, 2017; pp. 1–4.
50. Huang, H.; Lin, L.; Tong, R.; Hu, H.; Zhang, Q.; Iwamoto, Y.; Han, X.; Chen, Y.W.; Wu, J. Unet 3+: A full-scale connected unet for medical image segmentation. In Proceedings of the ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Virtual, 4–9 May 2020; IEEE: New York, NY, USA, 2020; pp. 1055–1059.
51. Silberman, N.; Hoiem, D.; Kohli, P.; Fergus, R. Indoor segmentation and support inference from RGBD images. In Proceedings of the Computer Vision—ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, 7–13 October 2012; Springer: Berlin/Heidelberg, Germany, 2012; pp. 746–760.
52. Li, B.; Ren, W.; Fu, D.; Tao, D.; Feng, D.; Zheng, W.; Wang, Y. Benchmarking Single-Image Dehazing and Beyond. *IEEE Trans. Image Process.* **2019**, *28*, 492–505. [[CrossRef](#)]
53. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
54. Galdran, A. Image dehazing by artificial multiple-exposure image fusion. *Signal Process.* **2018**, *149*, 135–147. [[CrossRef](#)]
55. Chen, D.; He, M.; Fan, Q.; Liao, J.; Zhang, L.; Hou, D.; Yuan, L.; Hua, G. Gated context aggregation network for image dehazing and deraining. In Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa Village, HI, USA, 7–11 January 2019; pp. 1375–1383.
56. Tian, Z.; He, T.; Shen, C.; Yan, Y. Decoders matter for semantic segmentation: Data-dependent decoding enables flexible feature aggregation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 3126–3135.

**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.