

Article

IFE-Net: An Integrated Feature Extraction Network for Single-Image Dehazing

Can Leng and Gang Liu *

College of Applied Mathematics, Chengdu University of Information Technology, Chengdu 610025, China; 18284557954@163.com

* Correspondence: liugang@cuit.edu.cn

Abstract: In recent years, numerous single-image dehazing algorithms have made significant progress; however, dehazing still presents a challenge, particularly in complex real-world scenarios. In fact, single-image dehazing is an inherently ill-posed problem, as scene transmission relies on unknown and nonhomogeneous depth information. This study proposes a novel end-to-end single-image dehazing method called the Integrated Feature Extraction Network (IFE-Net). Instead of estimating the transmission matrix and atmospheric light separately, IFE-Net directly generates the clean image using a lightweight CNN. During the dehazing process, texture details are often lost. To address this issue, an attention mechanism module is introduced in IFE-Net to handle different information impartially. Additionally, a new nonlinear activation function is proposed in IFE-Net, known as a bilateral constrained rectifier linear unit (BCReLU). Extensive experiments were conducted to evaluate the performance of IFE-Net. The results demonstrate that IFE-Net outperforms other single-image haze removal algorithms in terms of both PSNR and SSIM. In the SOTS dataset, IFE-Net achieves a PSNR value of 24.63 and an SSIM value of 0.905. In the ITS dataset, the PSNR value is 25.62, and the SSIM value reaches 0.925. The quantitative results of the synthesized images are either superior to or comparable with those obtained via other advanced algorithms. Moreover, IFE-Net also exhibits significant subjective visual quality advantages.

Keywords: BCReLU; feature extraction; convolutional neural network; single-image dehazing



Citation: Leng, C.; Liu, G. IFE-Net: An Integrated Feature Extraction Network for Single-Image Dehazing. *Appl. Sci.* **2023**, *13*, 12236. <https://doi.org/10.3390/app132212236>

Academic Editors: José Miguel Molina Martínez and Krzysztof Koszela

Received: 28 September 2023
Revised: 29 October 2023
Accepted: 8 November 2023
Published: 11 November 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

Obtaining a clear and haze-free image is crucial in photography and computer vision applications. Due to the presence of a large amount of dust, smoke, mist, or other floating particles in the atmosphere, when the camera captures images in this environment, significant quality degradation often occurs in the resulting images. These degradations, in turn, may have a negative impact on the performance of many computer vision systems [1–4], such as detection, tracking, and classification. Therefore, restoring clean images from damaged inputs through image dehazing is extremely important in the field of computer vision.

To overcome quality issues caused via haze in captured images, the atmospheric scattering model [5–7] has been proposed to restore clean images; it can be formally written as follows:

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

where $I(x)$ is the observed hazy image, $J(x)$ is the true scene radiance, α is the global atmospheric light, $t(x)$ is the medium transmission map, and x is the pixel index in the observed hazy image I . Furthermore, the medium transmission map can be expressed as follows:

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

where $d(x)$ is the distance from the scene point to the camera, and β represents the attenuation coefficient of the atmosphere.

From Equation (1), it can be seen that the dehazing process requires the accurate estimation of the transmission map and atmospheric light. A small portion of research mainly focuses on estimating atmospheric light [8–12], but the accuracy of the atmospheric light obtained will directly affect the results after dehazing and excessive errors will lead to a decrease in the dehazing performance on the image. Alternative other algorithms focus more on accurately estimating transmission maps, and the estimation of a transmission map mainly falls into two categories: prior-based methods [13,14] and learning-based methods [15,16]. In order to compensate for information loss during image processing, some methods use different priors to obtain atmospheric light and transmission maps. For example, Berman et al. [17] proposed a non-local prior-based dehazing algorithm based on the assumption that the colors of clean images are well approximated by different colors. Based on the difference in brightness, the saturation of blurred images is blurred, and color attenuation prior (CAP) [18] is proposed to estimate scene depth. The image prior obtained using prior-based algorithms can easily be inconsistent with practice, which may lead to incorrect transmission approximations. Learning-based methods are effective and superior to prior based-algorithms, exhibiting significant performance improvements. In [19], two subnetworks were designed to estimate the transmission map and atmospheric light, respectively. In [20], the authors created three different images from the hazy image and fused the results of the three images after dehazing. However, deep learning-based methods require training on a large number of real hazy images and their corresponding images without haze. The methods of estimating atmospheric light and transmission maps separately have made significant progress, but both have limitations. On the one hand, the inaccurate estimation of transmission maps may lead to low image quality; on the other hand, the separate estimation of atmospheric light and transmission maps leads to difficulties in finding the inherent relationship between them.

In order to find the intrinsic relationship between the parameters of Equation (1), Boyi Li et al. [21] first proposed a dehazing model that does not estimate α and $t(x)$. This model directly generates clean images from blurred images, rather than relying on any separate intermediate parameter estimation steps. Recently, many methods have used end-to-end learning instead of atmospheric scattering models to directly obtain clean images from networks [22–25]. Another widely used method tends to predict the residual of potential haze-free images or haze-free images relative to hazy images, as they often achieve better performance [26–30]. Although these recent dehazing methods have made significant progress, due to the complex haze distribution and the difficulty in collecting image pairs during the training process, it is easy to lose image details during the dehazing process using limited a dataset.

Due to the difficulty in collecting image pairs during the training process, IFE-Net uses end-to-end models, adaptively learns network features, and adopts multiscale feature extraction to better extract information. In addition, parallel convolutional layers of different sizes are used to extract features from input images of different scales [31,32]. This feature extraction structure is conducive to preserving more information and reducing the loss of image details.

Considering the potential cumulative error caused via the separate estimation of atmospheric light and the transmission map, IFE-Net unifies atmospheric light and transmission maps as one parameter to directly obtain a clean image. In addition, attention mechanisms have been widely applied in the design of neural networks [19,33–36], which can provide additional flexibility in the network. Inspired by these works and considering the different weights of features in different regions, a feature attention mechanism module called attention mechanism (AM) is designed in the network, which processes different types of information more effectively.

In deep learning networks, the activation function is a nonlinear function that enables neural networks to learn and represent complex patterns and relationships. The selection

of the final activation function has a significant impact on the output results of the model, as different activation functions have different characteristics and applicable scenarios. We considered that the output of the last layer of the image after dehazing should have upper and lower boundaries. In IFE-Net, we designed a new activation function called a bilateral constrained rectifier linear unit (BCReLU). The specific details of BCReLU and its comparison with other activation results in the network are detailed in Section 3.2.3.

The main contributions are as follows:

1. IFE-Net directly produces the clean image from a hazy image, rather than estimating the transmission map and atmospheric light separately. All parameters of IFE-Net are estimated in a unified model.
2. We propose a novel attention mechanism (AM) module, which consists of a channel attention mechanism, pixel attention mechanism, and texture attention. This module has different weighted information for different features and focuses more on strong features in areas with thick haze.
3. A bilateral constrained rectifier linear unit (BCReLU) is proposed in IFE-Net. To our knowledge, no one else has proposed BCReLU. Its significance in obtaining image restoration is demonstrated through experiments.
4. The experiments show that IFE-Net performs well both qualitatively and quantitatively. The extensive experimental results also illustrate the effectiveness of IFE-Net.

2. Related Work

Recently, single-image dehazing has attracted widespread attention in the field of computer vision. Due to the unknown global atmospheric light and transmission map, single-image dehazing is an inherently ill-posed problem. Many different methods have been proposed to address the issue. These methods can be roughly divided into prior-based and learning-based methods. The main difference between these two methods is that the prior-based methods mainly utilize prior statistical knowledge and hand-crafted features to process the hazy images, while the learning-based methods can automatically learn from the training set through a neural network and save it in the network's weights.

Single-image dehazing methods that extensively utilize prior knowledge have emerged. A patch-based dark channel prior (DCP) [11] method proposed by He et al. is one of the representative prior methods. Based on the assumption that hazy images may have extremely low intensity in at least one color channel, DCP uses an atmospheric scattering model for haze removal. Pixel-based dehazing methods [37,38] provide another solution to estimate the transmission map; however, pixel-based dehazing methods may result in insufficient information and an inability to estimate transmission maps. In addition, a method for establishing a linear model based on local prior images was proposed by Zhu et al. [18] to restore depth information. Although prior-based methods have achieved good results, the existence of priors is conditional. These hand-crafted priors are only applicable to specific cases and may not be applicable in changing scenarios.

The human brain is able to quickly distinguish hazy regions in natural images without other information, and convolutional neural networks have been inspired by this to be applied in image dehazing. These learning-based methods demonstrate extremely strong capabilities in dehazing. For example, Cai et al. [31] proposed Dehaze-Net, which is a trainable end-to-end network consisting of four parts: feature extraction; multiscale mapping; local extremum; and nonlinear regression. It is used to estimate the transmission map, and then the output transmission map is restored to a clean image through an atmospheric scattering model. Ren et al. [39] further proposed a multiscale convolutional neural network (MSCNN) for estimating scene transmission maps. Qin et al. [36] proposed an end-to-end feature fusion attention network (FFA-Net) to directly recover clean images, taking into account different weighted information. Due to the difficulty in obtaining paired clean images and hazy images in nature, Li et al. [40] studied the implementation of image dehazing without training on real clean image sets on the ground. These learning-

based methods have achieved good performance in dehazing and are more widely used in image dehazing.

3. The Proposed Method

In this section, we first introduce the transformed atmospheric scattering model. Then, a detailed introduction to the specific structure of the proposed IFE-Net is provided.

3.1. The Transformed Atmospheric Scattering Model

We can rewrite Equation (1) for the clean image as the output:

$$J(x) = \frac{I(x) - \alpha(1 - t(x))}{t(x)}. \quad (3)$$

Existing works such as [19,31,41] usually utilize empirical rules to estimate α and deep learning models to estimate $t(x)$. Estimating α and $t(x)$ separately will lead to certain errors. The output clean image obtained by combining α and $t(x)$ may have a greater cumulative error.

The transformed atmospheric scattering model is proposed [21] to reduce the cumulative error caused by separate estimating. The two parameters, α and $t(x)$, are unified into one formula to avoid the potential cumulative error caused by estimating the two parameters separately. Model (3) can be rewritten as follows:

$$J(x) = D(x)I(x) - D(x) + 1, \quad (4)$$

where

$$D(x) = \frac{\frac{I(x)-\alpha}{t(x)} + (\alpha - 1)}{I(x) - 1}. \quad (5)$$

It is worth noting that in Equation (5), α and $t(x)$ together form a new variable $D(x)$. The clean image can be obtained by estimating $D(x)$. The unified variable $D(x)$ can effectively reduce the cumulative error caused by estimating α and $t(x)$ separately.

3.2. Network Design

The architecture of the proposed IFE-Net contains three essential parts: (i) fused filters of different sizes concatenated them together to form a multiscale feature block; (ii) an attention mechanism composed of channel attention, pixel attention, and texture attention; and (iii) a bilateral constrained rectifier linear unit (BCReLU). As illustrated in Figure 1, the input image is first passed to the multiscale feature extraction block to produce multiscale features. Next, we process multiscale features using an attention mechanism block. The combination of the multiscale feature block and attention mechanisms module forms the $D(x)$ estimation block. Finally, we employ BCReLU to perform nonlinear regression on $D(x)$, thus obtaining the clean image.

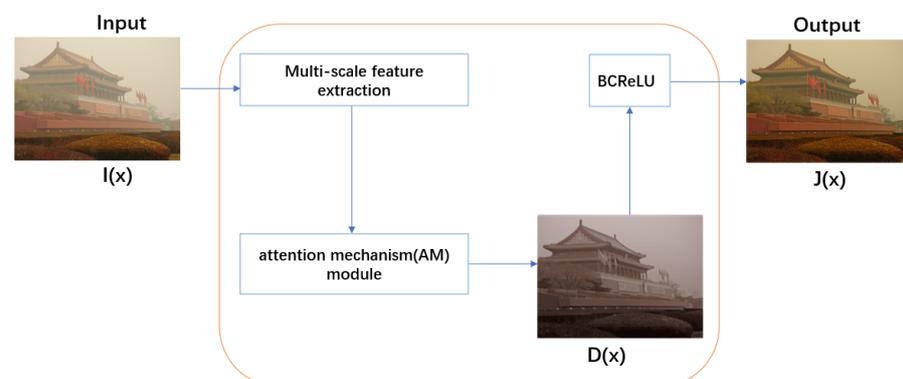


Figure 1. Overall architecture of IFE-Net.

3.2.1. Multiscale Feature Extraction

Multiscale feature extraction is very effective in the field of dehazing, while maintaining scale invariance and extracting information [42–45]. Moreover, parallel convolutions with different filter sizes are used to capture features at different scales. To compensate for the loss during information convolution, we connect network features of different scales to each other before extracting information from the next feature layer. Inspired by feature extraction methods, we used convolutional layers of different sizes to densely extract the features of the input image at different scales. As depicted in Figure 2, we choose to use five convolutional operations in the multiscale feature extraction block of IFE-Net, where the size of any convolution filter is among 1×1 , 3×3 , 5×5 , 7×7 , and 9×9 . “Conv1” uses a 1×1 convolution kernel to extract features, while “Conv2” uses a 3×3 convolution kernel to extract features; then, “Conv1” and “Conv2” layers are concatenated into the “concat1” layer. During the forward propagation process, a 5×5 sized convolution kernel is used to extract features from the “concat1” layer to obtain the “Conv3” layer. The “Conv2” layer, and the “Conv3” layer are concatenated into the “concat2” layer, and a 7×7 sized convolution kernel is used to extract features from the “concat2” layer to obtain the “Conv4” layer. Then, the “Conv3” layer and the “Conv4” layer are concatenated into the “Concat3” layer, and a 9×9 convolution kernel is used to extract features from the “concat3” layer to obtain the “Conv5” layer. Finally, the “Conv1” layer, “Conv2” layer, “Conv3” layer, “Conv4” layer and “Conv5” layer are concatenated to obtain the output of the multiscale feature feature extraction block. Importantly, the multiscale design of IFE-Net reduces information loss during convolutions and captures features at different scales.

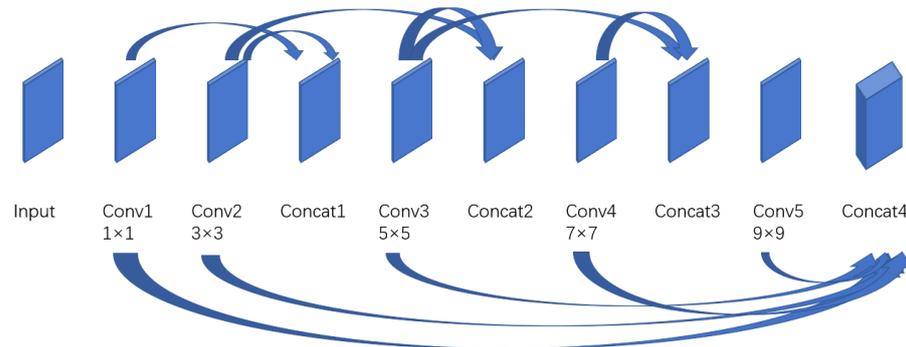


Figure 2. Multiscale feature extraction block.

3.2.2. Attention Mechanism

Most previous networks have treated channel and pixel features equally during image dehazing, resulting in unsatisfactory results after dehazing. Meanwhile, for images with uneven haze distribution, such networks cannot achieve good results. In order to better handle different parts of information, we designed a novel attention mechanism module, as shown in Figure 3. Compared to networks that treat channel and pixel features equally, the attention mechanism module of IFE-Net assigns different weights to different regions based on the importance of features. The more information the features contain, the greater their weight values. The application of the attention mechanism in IFE-Net focuses more on learning important information with high weights. The channel attention mechanism, pixel attention mechanism, and texture attention of the attention mechanism module can be expressed separately as follows:

$$F_1(y) = \text{ReLU}[\text{Conv}_7[\text{ReLU}(\text{Conv}_6(y))]], \quad (6)$$

$$F_2(y) = \text{ReLU}[\text{Conv}_9[\text{ReLU}(\text{Conv}_8(F_1(y)))]], \quad (7)$$

$$F_3(y) = \text{pool}(F_2(y)), \quad (8)$$

where y is the output of the multiscale feature extraction block, which serves as the input of the attention mechanism module; $\text{Conv}_6(y)$ and $\text{Conv}_7(y)$ denote the 1×1 convolution layer; $\text{Conv}_8(y)$ and $\text{Conv}_9(y)$ denote the 3×3 convolution layer; pool represents the 5×5 channel maxpooling operation; $F_1(y)$ denotes the output of the channel attention mechanism; $F_2(y)$ denotes the output of the pixel attention mechanism; and $F_3(y)$ denotes the output of texture attention. The attention mechanism block effectively assigns different weights to the features of different regions, enabling the entire network architecture to better retain effective information while suppressing the impact of unimportant information.

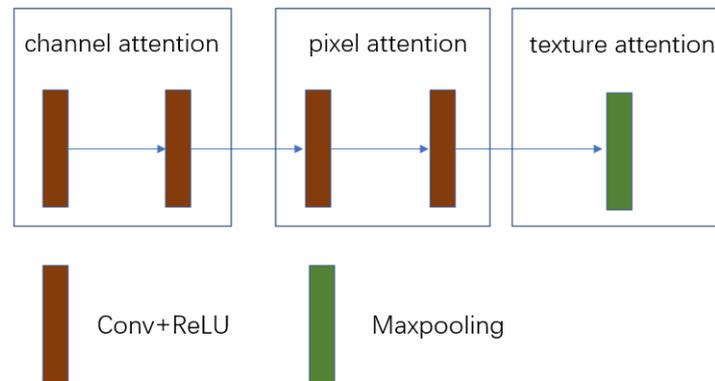


Figure 3. Attention mechanism module.

3.2.3. Bilateral Constrained Rectifier Linear Unit

Common choices for nonlinear activation function in deep networks include sigmoid [46], tanh [47], and rectified linear unit (ReLU) [48]. Sigmoid is a function of saturation at both ends, but it has a high computational cost and easily suffers from vanishing gradients, which may result in poor local optimality for the training process. Compared to sigmoid, tanh has an output mean of 0, which leads to faster convergence speed and fewer iterations. However, tanh, like sigmoid, has soft saturation, resulting in gradients vanishing. ReLU is proposed to alleviate the vanishing gradient problem of neural networks to a certain extent and accelerate the rate of convergence of gradient descent. It is worth noting that ReLU maintains unilateral suppression and has a wide area when it is greater than 0, which may lead to response overflow, especially in the final layer. For image restoration, the output of the last layer should have upper and lower boundaries, and the range of values should be relatively small. To this end, we propose a bilateral constrained rectifier linear unit (BCReLU) activation function to overcome the limitations of sigmoid and ReLU, as shown in Figure 4. As a novel linear unit, BCReLU keeps bilateral constraint and local linearity. Its output is centered around zero, making the latter layer of neurons less prone to bias and neuronal necrosis. In addition, BCReLU saves computational time and converges faster than other activation functions, which can help solve the gradient attenuation phenomenon as the number of layers increases. The marginal value of BCReLU is y_{max} and y_{min} ($y_{max} = 1$ and $y_{min} = -1$). BCReLU can be expressed as

$$\text{BCReLU}(x) = \begin{cases} 1, & x > 1 \\ x, & -1 \leq x \leq 1 \\ -1, & x < -1 \end{cases} \quad (9)$$

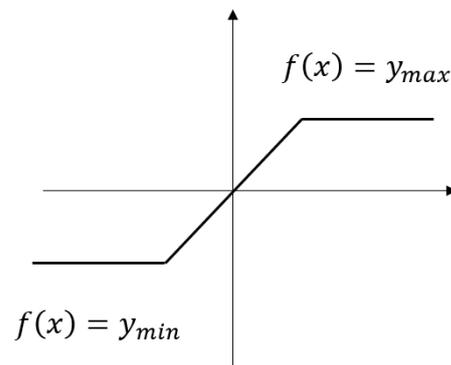


Figure 4. Bilateral constrained rectifier linear unit(BCReLU).

We compared the activation functions of the last layer in the network. Table 1 shows the quantitative evaluation results of different activation functions in the last layer on SOTS and ITS datasets (see Section 4.2 for details of PSNR and SSIM indicators). When using BCReLU in the last layer, the network achieved the best results, which confirms its effectiveness in IFE-Net.

Table 1. Quantitative results of quality evaluation indicators on SOTS and ITS datasets using different activation functions in the final layer.

Evaluation Indicators	ReLU	Tanh	Sigmoid	BCRelu
PSNR (SOTS)	24.59	20.07	18.61	24.63
SSIM (SOTS)	0.904	0.901	0.859	0.905
PSNR (ITS)	25.31	23.97	22.21	25.62
SSIM (ITS)	0.905	0.924	0.902	0.925

4. Experiments

To verify the superiority of IFE-Net, the dehazing results of IFE-Net were qualitatively and quantitatively compared with those of existing advanced dehazing methods using real-world images and benchmark datasets.

4.1. Datasets and Implementation Details

We chose the ground truth images with depth meta-data from the indoor NYU2 Depth Database [49]. Over 1440 clean images were selected from the NYU2 database and used to create synthesized hazy images using Equation (1). We chose $\beta \in \{0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6\}$, and each channel was set with different atmospheric light A , with a range of $[0.6, 1.0]$. The synthesized training set includes 27,193 haze images, and the learning rate is set to 0.0001 during the training process.

During the course of the experiments, we adopted the simple mean square error (MSE) loss function. Moreover, we utilized the BCReLU neuron in the last convolutional layer, as we find it more effective than other neurons, in our specific environment. The IFE-Net only needs a few epochs to converge and exhibits stability after approximately 10 epochs. In this study, we save the model parameters for 10 epochs of training for dehazing. We notice that an appropriately large batchsize can yield good performance in the batch normalization layer [50]. Due to limited physical memory on the GPU cards, the batch size of images is set to 16 during training. All experiments are performed on an NVIDIA RTX 3060 Ti GPU, NVIDIA, Santa Clara, CA, USA.

4.2. Quantitative Results on Synthetic Images

We adopt the peak signal-to-noise ratio (PSNR) and structure similarity index measure (SSIM) [51] as image quality indicators for quantitative analysis. PSNR is generally used to

measure the reference value of image quality between the maximum signal and background noise, and the larger the value, the lower the image distortion. PSNR can be expressed as

$$\text{PSNR} = 10 \times \log_{10}\left(\frac{(\text{MaxValue})^2}{\text{MSE}}\right), \quad (10)$$

where MSE is the mean square error of two images, and MaxValue is the maximum value that can be obtained from image pixels. SSIM is an indicator that measures the similarity between two images. From the perspective of image composition, SSIM defines structural information as independent of brightness and contrast, reflecting the properties of object structures in the scene. SSIM models distortion as a combination of three different factors: brightness, contrast, and structure. It uses the mean as the estimate of brightness, standard deviation as the estimate of contrast, and covariance as the measure of structural similarity.

$$\text{SSIM}(x, y) = \frac{(2u_x u_y + C_1)(2\sigma_{xy} + C_2)}{(u_x^2 + u_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (11)$$

where $C_1 = (K_1 L)^2$ and $C_2 = (K_2 L)^2$ are used to avoid situations where the denominator is 0; L is equivalent to MaxValue in PSNR, which is a very small constant; u_x and u_y are the mean; σ_x^2 is the standard deviation; and σ_{xy} is the variance. Compared to PSNR, SSIM is more in line with human visual characteristics in evaluating image quality.

High PSNR and SSIM scores indicate low image distortion and a more similar structure. We compare IFE-Net with the powerful methods in recent years based on PSNR and SSIM indicators. DCP [11] does not require precise physical modeling of haze in images but only relies on the prior principle of dark channels to reliably calculate the transmission matrix for image dehazing. Dehaze-Net [31] is an end-to-end system that utilizes prior knowledge to obtain atmospheric light, only learns the medium transmission map through the network, and ultimately obtains clean images. AOD-Net [21] is the first end-to-end trainable dehazing model, which does not separately estimate the parameters in Equation (1) but rather unifies all parameters into one and directly obtains a clean image from the hazy image. FFA-Net [36] has a feature fusion attention mechanism, and the design of the network allows it to perform well with dense hazes, textures, and details. GCA-Net [52] applies gated subnetworks and smooth extended convolutions, which is beneficial for fusing features of different scales and removing possible grid artifacts. DWGAN [53] introduces 2D discrete wavelet transform, aiming at restoring clear texture details and retaining sufficient high-frequency information. GUNet [54] significantly reduces overhead while effectively removing haze. The images in the RESIDE dataset [55] were selected for experimental evaluation of our method.

Figure 5 shows the dehazing results of some randomly selected synthetic images from the SOTS datasets. DCP [11], Dehaze-Net [14], and DWGAN [53] successfully remove heavy haze, but they exhibit color distortion and increased brightness. There are also issues with brightness enhancement and contrast in the results generated via FFA-Net [36], GCA-Net [52], GUNet [54], and AOD-Net [25]. IFE-Net handles details better and maintains color consistency with the ground truth. From the results, it can be observed that the results of IFE-Net are significantly better than other networks in terms of fidelity of image details and color. Table 2 shows the average quantitative results of the quality evaluation indicators in Figure 5, and the PSNR and SSIM values of IFE-Net are superior to the other methods. Tables 3 and 4 show the PSNR and SSIM results of our images after dehazing, respectively. Meanwhile, Table 5 shows the average time it takes for different networks to process each image with a size of 548×412 . The results in Tables 2–5 indicate that IFE-Net is effective and efficient.

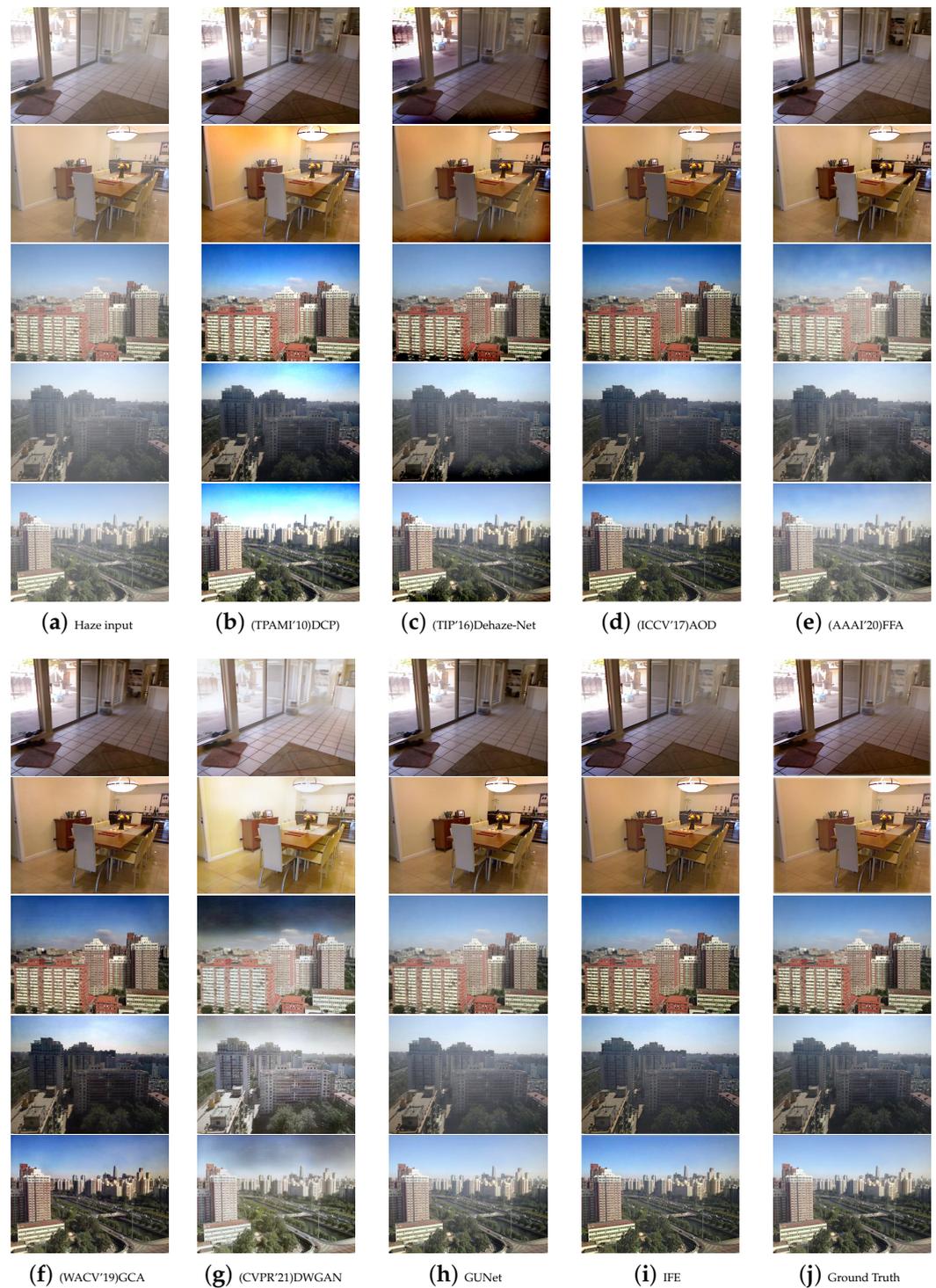


Figure 5. Quantitative comparison of IFE-Net with other methods on SOTS.

Table 2. The average quantitative results of the quality evaluation indicators in Figure 5.

Evaluation Indicators	DCP	Dehaze-Net	AOD	FFA	GCA	DWGAN	GUNet	IFE
PSNR	20.37	20.66	22.51	20.87	19.52	14.27	21.97	24.38
SSIM	0.913	0.886	0.928	0.909	0.902	0.815	0.921	0.942

Table 3. Quantitative results of quality evaluation indicators on SOTS dataset.

Evaluation Indicators	DCP	Dehaze-Net	AOD	FFA	GCA	DWGAN	GUNet	IFE
PSNR	21.37	21.34	22.12	21.31	23.05	20.56	19.382	24.63
SSIM	0.892	0.857	0.903	0.881	0.889	0.901	0.924	0.905

Table 4. Quantitative results of quality evaluation indicators on ITS dataset.

Evaluation Indicators	DCP	Dehaze-Net	AOD	FFA	GCA	DWGAN	GUNet	IFE
PSNR	20.32	18.71	22.39	18.48	27.77	14.79	19.26	25.62
SSIM	0.887	0.888	0.917	0.887	0.936	0.850	0.899	0.925

Table 5. The average time taken by different networks to process each image.

Metrics	DCP	Dehaze-Net	AOD	FFA	GCA	DWGAN	GUNet	IFE
Time (In seconds)	0.1294	0.6221	0.0194	0.6089	0.0592	0.1330	0.1106	0.0249

4.3. Qualitative Results on Real-World Images

Figure 6 shows a comparison of the results between IFE-Net and other methods using real scenes. As shown in Figure 6, DCP and GCA suffered from visual artifacts on the real hazy images. AOD-Net, FFA-Net, and GCA produced unrealistic colors in one or several images, such as the results of AOD-Net and FFA-Net in the fourth row as well as the results of AOD-Net and GCA in the fifth row. Dehaze-Net and FFA-Net retained a thin layer of haze, as shown in the second row. However, IFE-Net achieved excellent results in both thin and thick haze areas while maintaining colors consistent with real scenes. A similar result can be observed in the outdoor images shown in Figure 7. We enlarged the upper left corner of Figure 7a–g,h–n to show the enlarged results. The results of AOD-Net, FFA-Net, GCA, and GUNet exhibited color distortions and many non-natural characteristics. Additional white haze appeared in FFA-Net, resulting in incomplete dehazing. The result of DWGAN contained too much white. GCA and DWGAN showed unclear outlines of buildings in hazy areas above the sky. In contrast, IFE-Net successfully removed almost all of the haze while preserving the essential properties of the images, with obvious advantages in preserving edges, texture, contrast, brightness, and other image characteristics, as shown in Figure 7.

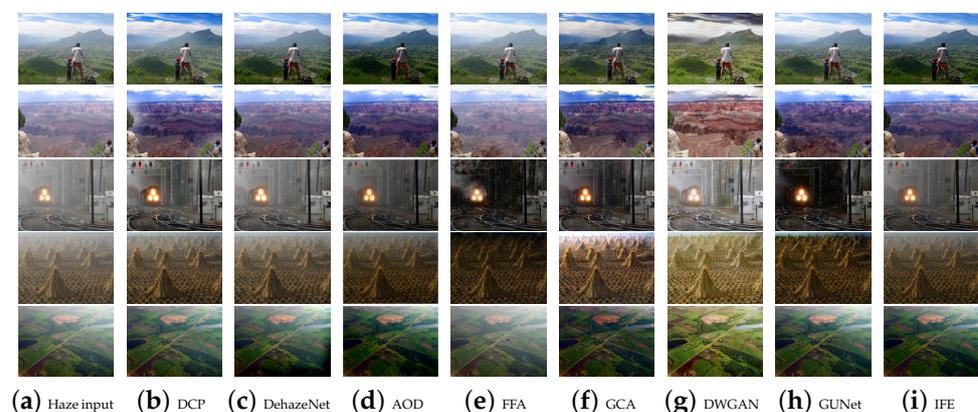


Figure 6. Qualitative comparisons of IFE-Net on real-world images.

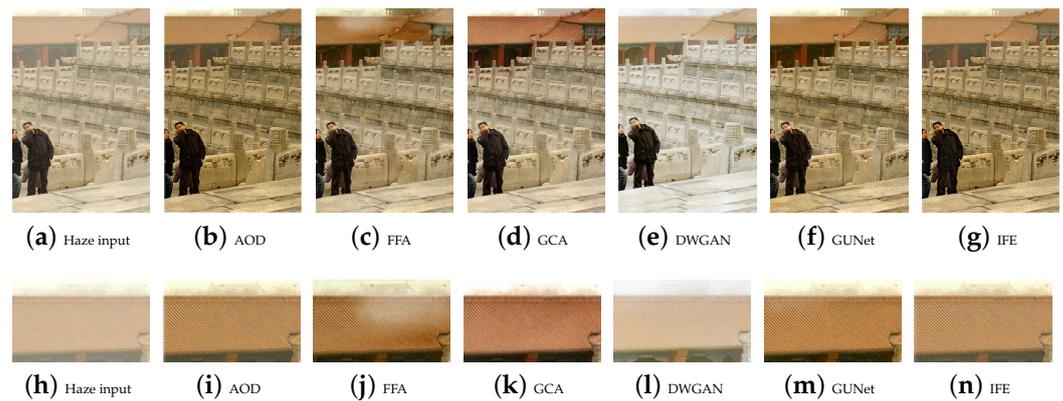


Figure 7. This figure shows our ability in detail and color processing compared to other networks.

In addition, we also removed haze from the hazy image of a large area of the sky and compared it with several other advanced methods. Most dehazing algorithms have poor dehazing effects on images containing large areas of sky, resulting in color distortion and uneven color blocks in the restored haze-free images. We show the results of several methods in Figure 8. Figure 8a shows the input hazy image, and Figure 8b–d show the dehazing results of GCA, GUN-Net, and IFE-Net, respectively. From Figure 8b, it can be observed that the results obtained show a thorough removal of haze on the ground, but uneven color blocks appear on the ground and also in the sky area. In Figure 8c, there are no issues with image color distortion, but the dehazing effect in the ground and sky areas is not significant. In Figure 8d, the haze in the sky is suppressed without significant color distortion blocks. Simultaneously, the dehazing effect in the ground area is significant, and the results obtained are good in terms of dehazing and details.

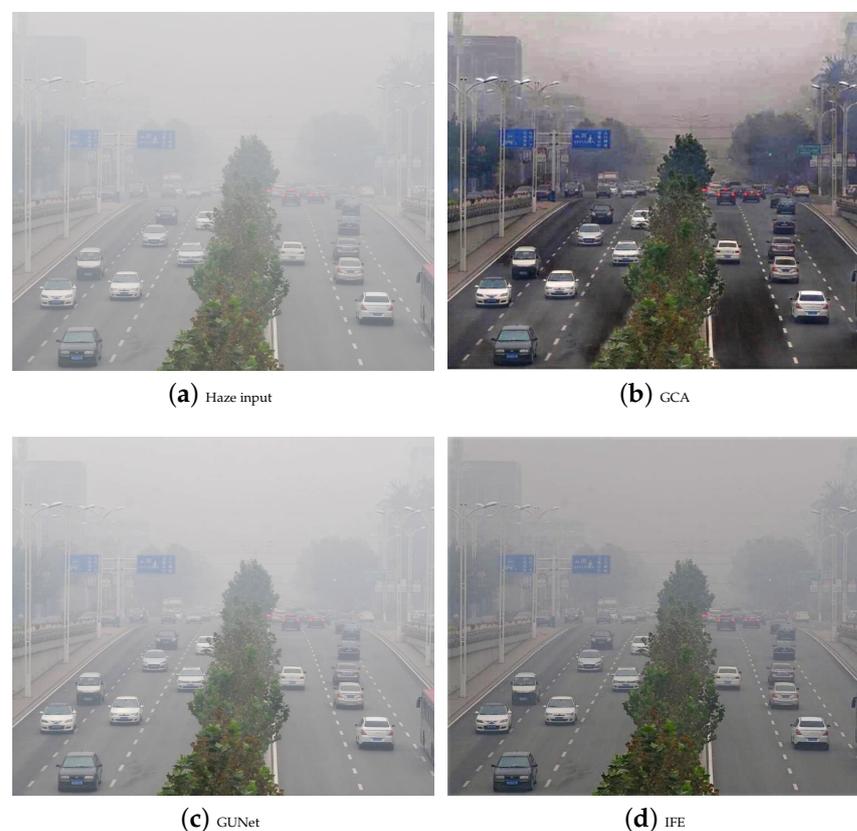


Figure 8. Results of dehazing in sky areas with dense haze.

4.4. Ablation Research

Both IFE-Net and AOD unify the atmospheric light and transmission map in the atmospheric scattering model into one parameter, directly obtaining clean images. In order to evaluate the contribution of the AM module in the network, we compared the networks with and without it in AOD and IFE-Net, respectively. Table 6 shows the experimental results on two datasets, indicating that the addition of AM modules resulted in better PSNR and SSIM results. Figure 9 shows a comparison of the visual effects of images; networks without an AM module have darker colors, while networks with an AM module achieve better visual effects. The quantitative and qualitative results in the ablation research demonstrate the effectiveness of an AM module in the networks.

Table 6. Effectiveness of AM module.

Dataset	SOTS		ITS	
	PSNR	SSIM	PSNR	SSIM
AOD	22.12	0.903	22.39	0.917
AOD + AM	24.13	0.904	23.99	0.920
IFE without AM	23.16	0.902	23.77	0.921
IFE + AM	24.63	0.905	25.62	0.925



Figure 9. Comparison of the visual effects of AM block.

5. Conclusions

We proposed a novel end-to-end adaptive enhancement dehazing network, called IFE-Net, to address the challenge of single-image dehazing. IFE-Net consists of a multiscale feature extraction block, an attention mechanism (AM) module, and a bilateral constrained rectifier linear unit (BCReLU). Considering the cumulative errors that may arise from estimating atmospheric light and transmission maps separately, IFE-Net estimates a parameter that is unified by both. Its novel network design effectively performs feature extraction. In addition, we designed an attention mechanism (AM) module to address the varying importance of information in different regions. The importance of BCReLU in image restoration was also demonstrated through experiments. We compared IFE-Net with other dehazing methods using PSNR and SSIM, and the results show that IFE-Net achieved good scores for both indicators. At the same time, we used subjective criteria to analyze the results obtained via different methods on natural hazy images. Our conclusion

is that the proposed IFE-Net combines feature extraction blocks, attention mechanism, and a BCRReLU activation function, making it significantly effective in natural and synthetic image dehazing. Although our IFE-Net has a simple structure, it shows strong capabilities in haze removal. The experimental results confirm the superiority and efficiency of IFE-Net. At present, IFE-Net has achieved good results in dehazing, and another promising area for our future research is to apply it to image enhancement algorithms.

Author Contributions: Methodology, C.L.; Software, C.L.; Writing—original draft, C.L.; Writing—review & editing, G.L.; Supervision, G.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by National Nature Science Foundation of China (NSFC) 11901065 and Chengdu University of Information Technology (KYTD202243).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The source data NYU2 Depth Database are openly available at [https://doi.org/10.1007/978-3-642-33715-4_54], reference number [49]. The source data RESIDE dataset are openly available at [<https://doi.org/10.1109/TIP.2018.2867951>], reference number [55]. All the extension data are generated by ourselves and freely available to other researchers.

Acknowledgments: We thank the anonymous reviewers for their valuable comments.

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

References

1. Cucchiara, R.; Grana, C.; Piccardi, M.; Prati, A. Detecting moving objects, ghosts, and shadows in video streams. *IEEE Trans. Pattern Anal. Mach. Intell.* **2003**, *25*, 1337–1342. [[CrossRef](#)]
2. Jung, C.R. Efficient background subtraction and shadow removal for monochromatic video sequences. *IEEE Trans. Multimed.* **2009**, *11*, 571–577. [[CrossRef](#)]
3. Sanin, A.; Sanderson, C.; Lovell, B.C. Improved shadow removal for robust person tracking in surveillance scenarios. In Proceedings of the International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; pp. 141–144.
4. Zhang, W.; Zhao, X.; Morvan, J.; Chen, L. Improving shadow suppression for illumination robust face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *41*, 611–624. [[CrossRef](#)] [[PubMed](#)]
5. Cartney, E.J. *Optics of the Atmosphere: Scattering by Molecules and Particles*; John Wiley and Sons, Inc.: New York, NY, USA, 1976; 421p.
6. Narasimhan, S.G.; Nayar, S.K. Chromatic framework for vision in bad weather. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2000 (Cat. No. PR00662), Hilton Head, SC, USA, 13–15 June 2000; Volume 1, pp. 598–605.
7. Narasimhan, S.G.; Nayar, S.K. Vision and the atmosphere. *Int. J. Comput. Vis.* **2002**, *48*, 233–254. [[CrossRef](#)]
8. Sulami, M.; Glatzer, I.; Fattal, R.; Werman, M. Automatic recovery of the atmospheric light in hazy images. In Proceedings of the Computational Photography (ICCP), Santa Clara, CA, USA, 2–4 May 2014; pp. 1–11.
9. Berman, D.; Treibitz, T.; Avidan, S. Air-light estimation using haze-lines. In Proceedings of the Computational Photography (ICCP), Evanston, IL, USA, 13–15 May 2017; pp. 1–9.
10. Vidyamol, K.; Prakash, M.S. An Improved Dark Channel Prior for Fast Dehazing of Outdoor Images. In Proceedings of the 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 3–5 October 2022; pp. 1–6.
11. He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. PAMI* **2011**, *33*, 2341–2353.
12. Raikwar, S.; Tapaswi, S. Accurate and Robust Atmospheric Light Estimation for Single Image Dehazing. In Proceedings of the 2020 IEEE 7th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), Prayagraj, India, 27–29 November 2020; pp. 1–4.
13. Qasim, M.; Raja, G. SPIDE-Net: Spectral Prior-Based Image Dehazing and Enhancement Network. *IEEE Access* **2022**, *10*, 120296–120311. [[CrossRef](#)]
14. Ajith, A.P.; Vidyamol, K.; Devassy, B.R.; Manju, P. Dark Channel Prior based Single Image Dehazing of Daylight Captures. In Proceedings of the 2023 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA), Ernakulam, India, 20–21 January 2023; pp. 1–6.
15. Sharma, T.; Nalla, B.T.; Verma, N.K.; Vasikarla, S. FR-HDNet: Faster RCNN based Haze Detection Network for Image Dehazing. In Proceedings of the 2022 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington, DC, USA, 11–13 October 2022; pp. 1–8.

16. Zhang, H.; Sindagi, V.; Patel, V.M. Joint transmission map estimation and dehazing using deep networks. *arXiv* **2017**, arXiv:1708.00581.
17. Berman, D.; Treibitz, T.; Avidan, S. Non-local image dehazing. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 21–23 June 2016; pp. 1674–1682.
18. 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.
19. Zhang, H.; Patel, M.V. Densely connected pyramid dehazing network. *IEEE Conf. Comput. Vis. Pattern Recognit.* **2008**, *2008*, 3194–3203.
20. Ren, W.; Ma, L.; Zhang, J.; Pan, J.; Cao, X.; Liu, W.; Yang, M.H. Gated fusion network for single image dehazing. *arXiv* **2018**, arXiv:1804.00213.
21. Li, B.; Peng, X.; Wang, Z.; Xu, J.; Feng, D. An all-in-one network for dehazing and beyond. *arXiv* **2017**, arXiv:1707.06543.
22. Laha, S.; Foroosh, H. Haar Wavelet-Based Attention Network for Image Dehazing. In Proceedings of the 2022 IEEE International Conference on Image Processing (ICIP), Bordeaux, France, 16–19 October 2022; pp. 3948–3952.
23. Raj, N.B.; Venketeswaran, N. Single Image Haze Removal using a Generative Adversarial Network. In Proceedings of the International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), Chennai, India, 4–6 August 2018.
24. Parihar, A.S.; Singh, K.; Ganotra, A.; Yadav, A. Contrast Aware Image Dehazing using Generative Adversarial Network. In Proceedings of the 2022 2nd International Conference on Intelligent Technologies (CONIT), Hubli, India, 25–27 June 2022; pp. 1–6.
25. Bai, H.; Pan, J.; Xiang, X.; Tang, J. Self-guided image dehazing using progressive feature fusion. *IEEE Trans. Image Process.* **2022**, *31*, 1217–1229. [[CrossRef](#)] [[PubMed](#)]
26. Liu, X.; Ma, Y.; Shi, Z.; Chen, J. Grid dehazenet: Attention-based multi-scale network for image dehazing. *ICCV* **2019**, *2019*, 7314–7323.
27. Dong, J.; Pan, J. *Physics-Based Feature Dehazing Networks*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 188–204.
28. Deng, Q.; Huang, Z.; Tsai, C.; Lin, C. *Hardgan: A haze-Aware Representation Distillation Gan for Single Image Dehazing*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 722–738.
29. 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, 20–21 June 2021; pp. 10551–10560.
30. Wang, C.; Shen, H.; Fan, F.; Shao, M.; Yang, C.; Luo, J.; Deng, L. Eaa-net: A novel edge assisted attention network for single image dehazing. *KBS* **2021**, *228*, 107279. [[CrossRef](#)]
31. 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](#)]
32. Agrawal, S.C.; Jalal, A.S. Linear Fusion of Multi-Scale Transmissions for Image Dehazing. In Proceedings of the 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 6–8 July 2021; pp. 1–6.
33. Ye, F.; Wu, K.; Zhang, R.; Wang, M.; Meng, X.; Li, D. Multi-Scale Feature Fusion Based on PVTv2 for Deep Hash Remote Sensing Image Retrieval. *Remote Sens.* **2023**, *15*, 4729. [[CrossRef](#)]
34. Zong, P.; Li, J.; Hua, Z. Lightweight Multi-scale Attentional Network for Single Image Dehazing. In Proceedings of the 2022 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML), Xi'an, China, 28–30 October 2022; pp. 401–405.
35. Shit, S.; Das, D.K.; Sur, A.; Ray, D.N.; Banik, B.C.; Rana, A. Encoder and Decoder-Based Feature Fusion Network for Single Image Dehazing. In Proceedings of the 2023 3rd International Conference on Artificial Intelligence and Signal Processing (AISP), Vijayawada, India, 18–20 March 2023; pp. 1–5.
36. Qin, X.; Wang, Z.; Bai, Y.; Xie, X.; Jia, H. FFA-Net: Feature Fusion Attention Network for Single Image Dehazing. *Proc. AAAI Conf. Artif. Intell.* **2020**, *34*, 11908–11915. [[CrossRef](#)]
37. Tarel, J.P.; Hautiere, N. Fast visibility restoration from a single color or gray level image. In Proceedings of the IEEE 12th International Conference on Computer Vision, 29 September–2 October 2009; pp. 2201–2208.
38. Wang, W.; Yuan, X.; Wu, X.; Liu, Y. Fast image dehazing method based on linear transformation. *IEEE Trans. Multimedia* **2017**, *19*, 1142–1155. [[CrossRef](#)]
39. 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; pp. 154–169.
40. Li, B.; Gou, Y.; Gu, S.; Liu, J.Z.; Zhou, J.T.; Peng, X. You only look yourself: Unsupervised and untrained single image dehazing neural network. *Int. J. Comput. Vis.* **2021**, *129*, 1754–1767. [[CrossRef](#)]
41. Liu, J.; Liu, W.; Sun, J.; Zeng, T. Rank-one prior: Toward real-time scene recovery. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Online, 19–25 April 2021; pp. 14802–14810.
42. Purkayastha, P.; Choudhary, M.S.; Kumar, M. Steerable Pyramid-based Multi-Scale Fusion Algorithm for Single Image Dehazing. In Proceedings of the 2023 International Conference on Device Intelligence, Computing and Communication Technologies, (DICCT), Dehradun, India, 17–18 March 2023; pp. 552–557.

43. Zhao, L.; Zhang, Y.; Cui, Y. An attention encoder-decoder network based on generative adversarial network for remote sensing image dehazing. *IEEE Sensors J.* **2022**, *22*, 10890–10900. [[CrossRef](#)]
44. Zhong, T.; Cheng, M.; Dong, X.; Wu, N. Seismic random noise attenuation by applying multiscale denoising convolutional neural network. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–13. [[CrossRef](#)]
45. Zhang, J.; Shao, M.; Wan, Z.; Li, Y. Multi-scale feature mapping network for hyperspectral image super-resolution. *Remote. Sens.* **2021**, *13*, 4180. [[CrossRef](#)]
46. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 5786. [[CrossRef](#)]
47. LeCun, Y.; Bottou, L.; Orr, G.B.; Müller, K. *Efficient Backprop*; Springer: Berlin/Heidelberg, Germany, 1998; pp. 9–48.
48. Glorot, X.; Bordes, A.; Bengio, Y. Deep sparse rectifier neural networks. In Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, Ft. Lauderdale, FL, USA, 11–13 April 2011; pp. 315–323.
49. Silberman, N.; Hoiem, D.; Kohli, P.; Fergus, R. Indoor segmentation and support inference from rgb-d images. In Proceedings of the European Conference on Computer Vision, Florence, Italy, 7–13 October 2012; pp. 746–760.
50. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on Machine Learning, Lille, France, 7–9 July 2015.
51. 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](#)] [[PubMed](#)]
52. Chen, D.; He, M.; Fan, Q.; Liao, J.; Zhang, L.; Hou, D.; uan, L.Y.; Hua, G. Gated context aggregation network for image dehazing and deraining. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision, (WACV), Waikoloa Village, HI, USA, 7–11 January 2019; pp. 1375–1383.
53. Fu, M.; Liu, H.; Yu, Y.; Chen, J.; Wang, K. DW-GAN: A Discrete Wavelet Transform GAN for NonHomogeneous Dehazing. *arXiv* **2021**. [[CrossRef](#)]
54. Song, Y.; Zhou, Y.; Qian, H.; Du, X. Rethinking Performance Gains in Image Dehazing Networks. *arXiv* **2022**, arXiv:2209.11448.
55. Li, B.; Ren, W.; Fu, D.; Tao, D.; Feng, D.; Zeng, W.; Wang, Z. Benchmarking Single-Image Dehazing and Beyond. *IEEE Trans. Image Process.* **2019**, *28*, 492–505. [[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.