Article

A Comprehensive Review of Traditional and Deep-Learning-Based Defogging Algorithms

Minxian Shen [†] [ID]**, Tianyi Lv** [†] [ID]**, Yi Liu** [ID]**, Jialiang Zhang** [ID] **and Mingye Ju** *[ID]

School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing 210046, China; b23080421@njupt.edu.cn (M.S.); q23010312@njupt.edu.cn (T.L.); q23010116@njupt.edu.cn (Y.L.); b20031620@njupt.edu.cn (J.Z.)

* Correspondence: jumingye@njupt.edu.cn
† These authors contributed equally to this work.

Abstract: Images captured under adverse weather conditions often suffer from blurred textures and muted colors, which can impair the extraction of reliable information. Image defogging has emerged as a critical solution in computer vision to enhance the visual quality of such foggy images. However, there remains a lack of comprehensive studies that consolidate both traditional algorithm-based and deep learning-based defogging techniques. This paper presents a comprehensive survey of the currently proposed defogging techniques. Specifically, we first provide a fundamental classification of defogging methods: traditional techniques (including image enhancement approaches and physical-model-based defogging) and deep learning algorithms (such as network-based models and training strategy-based models). We then delve into a detailed discussion of each classification, introducing several representative image fog removal methods. Finally, we summarize their underlying principles, advantages, disadvantages, and give the prospects for future development.

Keywords: dehaze; image enhancement; traditional defogging algorithm; atmospheric scattering model; dark channel prior; CNN; unsupervised learning; deep learning

1. Introduction

Haze or fog in bad weather, which may result in dull colors and blurred contrast for captured images, is a traditional atmospheric phenomenon. Image dehazing is able to help to improve the clarity and visibility of such degraded images, which makes it widely used in unmanned, security monitoring, aerospace, and other fields to enhance the performance and safety of these fields. Since 2004, the publication of image defogging or dehazing has been increasing, and a large number of new methods have been proposed.

The essence of image defogging is to get rid of adverse visual effects caused by haze or fog. The most intuitive way to address such issue is to employ the traditional image enhancement techniques [1–8] to improve the contrast and saturation of fog images. However, these methods do not work well on fog images. This is due to the fact that these methods do not consider the mechanism that the deterioration of haze images is related to the haze concentration. To alleviate this problem, numerous atmospheric scattering model (ASM)-based [9] image dehazing techniques [10–21] have been developed. These methods are mainly based on prior knowledge, such as dark channel prior [10] and color attenuation prior [12], to reduce the uncertainty of ASM. Then, they use the estimated parameters to reversely restore the high-quality scene from single image.

Recently, with the sharp development of artificial intelligence (AI), by leveraging AI frameworks, many deep learning-based image dehazing methods have been further proposed. Compared to physical-model-based dehazing method, this type of algorithm can achieve better results because of its powerful fitting ability. In general, these image dehazing methods either make use of different types of network architectures or employ different loss functions during training. Therefore, they can be roughly divided by network used

type and training strategy. In the category of training methods, according to whether there is supervision information, it can be further divided into three categories: unsupervised image dehazing methods [22–30], supervised image dehazing methods [31–36], and semi-supervised image dehazing methods [37,38].

Although great progress in image dehazing or defogging have been made, there are few survey papers on such field, to the best of our knowledge. Liu et al. [39] provide a detailed summary of classical defogging methods, including depth estimation, wavelets, enhancement, and filtering, but lack a detailed discussion of the latest neural network models. Xu et al. [40] highlights image recovery algorithms, contrast enhancement algorithms, and fusion-based defogging algorithms. Additionally, they describe current video-defogging algorithms while still ignores an introduction to deep learning-based defogging algorithms. Although Gui et al. [41] give a full summary and discussion of neural networks and loss functions, they still lack a discussion of traditional defogging algorithms and a systematic classification for defogging methods. Ancuti et al. [42] carries on with some supervised defogging models. However, they do not focus on recent applications of unsupervised methods.

In this paper, we conduct a comprehensive overview of image dehazing or defogging techniques. Unlike the aforementioned survey works, this review firstly categorizes the classic or state-of-the-art image fog/haze removal algorithms in detail, including traditional image enhancement approaches, physical-model-based defogging, network-based dehazing models, and training strategy-based models. Moreover, we also conduct qualitative and quantitative comparisons of each type of algorithm, aiming at pointing out their advantages and disadvantages and raising the outlook that may boost the image defogging field.

The remainder of this paper is organized as follows. Following the introduction, Sections 2 and 3 introduce non-deep learning defogging and deep learning defogging, respectively. In detail, Section 2 depicts traditional image defogging algorithm and physical-model-based defogging, while Section 3 illustrates the architecture of the deep network and training strategy used for image defogging. Section 4 conducts extensive performance evaluations of the state-of-the-art approaches mentioned in above, and illustrates the advantages and disadvantages of each type of algorithm. Finally, the conclusions and outlook for future are drawn in Section 5.

2. Non-Deep Learning Defogging

Early image defogging either simply enhances the local or global contrast, or explores the hand-crafted priors on foggy images to achieve image fog removal. The former, namely traditional image defogging algorithm, directly exploits the traditional methods, e.g., histogram equalization, signal analysis methods, and other traditional contrast enhancement approaches. The latter, namely physical-model-based defogging, is based on a physical model to estimate the imaging parameters, thereby realizing high-quality fog removal.

2.1. Traditional Image Defogging Algorithm

2.1.1. Histogram Equalization

The core idea of histogram equalization (HE) is based on redistributing images' intensity levels. This method is conducted by calculating the histogram of the image and then generating the cumulative distribution function (CDF). Subsequently, a uniform distribution is obtained by mapping the original pixel values to new values according to the CDF. Due to the implementation of the aforementioned measures, this method effectively spreads out the pixel values across the intensity spectrum, thereby enhancing the visibility of image details. Considering the mechanism of histogram equalization, there exists a relationship between gray level and pixels:

$$s_k = T(r_k) = \sum_{i=0}^{k} p_r(r_i) = \sum_{i=0}^{k} \frac{n_i}{N} \tag{1}$$

where N is the total amount of pixels in the image, L represents the total number of gray levels, s_k is the value of the cumulative distribution function corresponding to the gray level r_k, $p_r(r_j)$ is the probability density function value of the gray level r_j in the original image, and T is the transform function [1].

Although HE is able to improve the global contrast of an image, it lacks a selection mechanism for processed signals, thus making it amplify noise. Subsequently, adaptive HE (AHE) was proposed, and it partitions images into blocks and then applies specific defogging methods to each block [2,3]. However, it brings a new issue, i.e., over-load computational complexity that may reduce its real-time performance. Contrast limited AHE (CLAHE), an improved version of AHE, which uses a bi-linear interpolation [4] to mitigates these issues. Unfortunately, CLAHE still exhibit noticeable block artifacts and substantial changes in overall brightness.

2.1.2. Signal Analysis-Based Approach

The most representative method of this type is the well-known homomorphic filtering. In homomorphic filtering, the input image is decomposed into two components. The first component represents the spatially varying incident light intensity, which changes slowly and is primarily present in the low-frequency regions of the foggy image. The second component encapsulates the scenario reflection perceived by the human eye, highlighting the intricacies and details of the scenario. To balance these two components, a logarithmic transformation was designed, whose key idea is to suppress low frequency and high frequency [5]. Mathematically, homomorphic filtering can be formulated as follows:

$$g(x,y) = H(u,v)F(u,v)$$
$$H(u,v) = \gamma(u,v) \cdot F_{lp}(u,v) + (1 - \gamma(u,v)) \cdot F_{hp}(u,v)$$

(2)

where $g(x,y)$ is the pixel value of the output image, $F(u,v)$ is the Fourier transform of the input image in the frequency domain, $H(u,v)$ is the frequency response function of the homomorphic filter, (u,v) is a variable in the frequency domain, and $\gamma(u,v)$ is a function that controls the degree of mixing of the low and high-frequency components.

$F_{lp}(u,v)$ and $F_{hp}(u,v)$ are the low-pass and high-pass filters. The advantages of homomorphics lie in the removal of multiplicative noise, enhancement of contrast in adjacent regions, and compression of the overall dynamic range of the image. However, there still exist several notable disadvantages in homomorphic filtering, e.g., being unable to deal with scenarios with severe fog conditions and lacking the ability to maintain the local details during processing. Therefore, to address such limitations, the Fourier transformation is replaced by the wavelet transform in homomorphic filtering for high-quality fog removal. Compared to the Fourier transformation, the wavelet transform provides a multi-dimensional transformation connecting spatial, temporal, and frequency domains. By incorporating the localized adaptation nature of short-time Fourier transformation and utilizing finite-length wavelets that decay, the wavelet transforms significantly enhance its capacity to process non-stationary signals [6].

$$WT(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) * \Psi(\frac{t - \tau}{a})$$

(3)

where the parameter a controls the wavelet us contraction scale, while t controls its translation. The wavelet transform outperforms homomorphic filtering in handling with local image details and enriching overall information, but it may alter image brightness, causing distortions like blurred edges. To address these, the two-dimensional wavelet transform and threshold function are introduced. This combination separates high- and low-frequency components, emphasizing or eliminating detail levels to enhance useful information. While it improves contrast and information in defogged images, it does not fully and effectively address image distortion and edge blurring.

2.1.3. Other Traditional Image Enhancement Used for Image Defogging

In addition to the aforementioned image enhancement algorithms, there are other competitive alternatives that utilize Partial Differential Equations (PDEs) [7]. These PDE-based methods integrate Laplace operators with Retinex algorithms. A notable advantage of PDE-based approaches is to offer a more physically explanatory and rational method compared to alternative techniques. By leveraging models of light propagation and scattering, PDE-based algorithms achieve a globally consistent fog removal effect. However, these methods also have certain disadvantages. In contrast to the histogram equalization and signal analysis-based defogging algorithms mentioned earlier, PDE-based methods exhibit higher computational complexity. Moreover, when handling with complex scenarios, the defogging process may cause overcompensation, which will lead to areas to be excessively bright or dark.

The Laplace operator-based algorithm aims to enhance image contrast through quadratic differentiation for sharpening. The main advantage of this method is its low complexity. However, it proves less effective in challenging defogging scenarios, often resulting in severe noise and artifacts. Another approach, the Retinex algorithm, enhances images by separating them into albedo and light components, thereby emphasizing details and contrast to achieve fog removal. Various implementations of Retinex algorithms include single-scale Retinex (SSR) [8] and multi-scale Retinex (MSR) [4]. MSR applies a convolution kernel at multiple scales, convolving the image to generate reflective images across these scales. These images are subsequently fused, weighted, and averaged to produce the final defogging output. While the multi-scale Retinex algorithm enhances defogging capabilities for diverse scenes and objects compared to the single-scale approach, it also significantly escalates computational complexity and introduces potential noise and artifacts.

While those algorithms employing image enhancement techniques have shown some effectiveness in defogging, their results remain room for improvement. Consequently, researchers have tried to introduce imaging model or deep learning theory to gain a better restoration performance.

2.2. Physical-Model-Based Defogging

This type of algorithm belongs to non-deep learning, and its essence is to search imaging parameters that are used for fog removal. Figure 1 showcases the schematic diagram for physical-model-based defogging. As shown, hand-crafted prior knowledge, e.g., dark channel prior (DCP), color attenuation prior (CAP), and gamma correction prior (GCP), is initially imposed on the atmospheric scattering model (ASM) [9] to derive the transmission map t and atmospheric light A. Then, these estimated parameters, along with the original foggy image, are fed into ASM to obtain the haze-free scene.
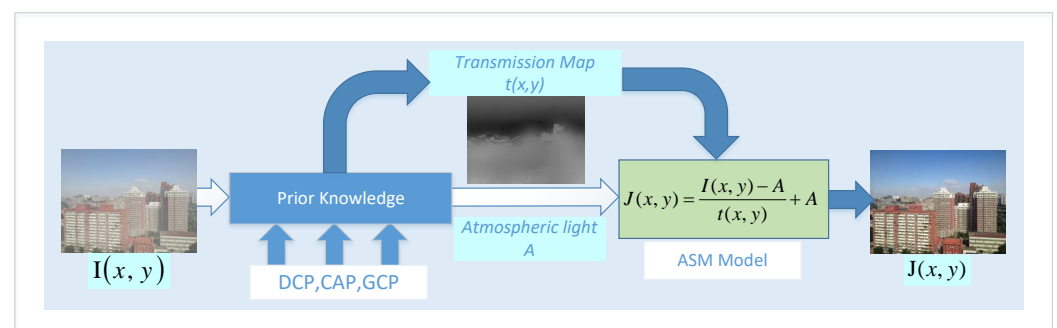

Figure 1. Schematic diagram for physical-model-based defogging approaches.

2.2.1. Atmospheric Scattering Model (ASM)

Before describing physical-model-based defogging approaches, it is necessary to introduce well-known ASM. Formally, the ASM can be further expressed by

$$I(x,y) = J(x,y)t(x,y) + A(1 - t(x,y)) \tag{4}$$

where A is the global atmospheric light, I is the haze image, J is the haze-free scene that is expected to be restored, and t is the transmission. In detail, the transmission can be written as

$$t = e^{-kd(x,y)} \tag{5}$$

where k and $d(x,y)$ represent the atmospheric scattering coefficient and scene depth, respectively. It is obvious from this equation that, on fog-free scenarios, k is very close to 0, which leads to $J = I$ according to ASM. When taking pictures on foggy scenarios, $k > 0$ cannot be ignored, and the light source received by the detector would be interfered with by fog. In this case, the collected light primarily originates from two sources: one is the target-reflected light attenuated by particles and detected by the system, while the other is atmospheric light resulting from particle scattering of the light source. Once these two parameters have been determined, the haze-free scenes can be easily restored by:

$$J(x,y) = \frac{I(x,y) - A}{t(x,y)} + A \tag{6}$$

where a useful solution of getting A and t is to impose prior knowledge or extra information on ASM. In the following, several classic defogging algorithms will be briefly outlined.

2.2.2. Dark Channel Prior Image Dehazing

The dark channel prior (DCP) defogging algorithm, introduced by He et al. [10], is known for its superior performance compared to other prior-based defogging methods. He et al. note that every region in a fog-free image contains at least bright colors or dark elements, such as shadows. Therefore, each region typically shows a channel with very low pixel values, known as the DCP. Formally, it can be defined as follows:

$$J^{\text{dark}}(x,y) = \min_{C \in \{R,G,B\}} \left(\min_{z \in \mu(x,y)} J^C(x,y) \right) \tag{7}$$

where c is the color index, $\mu(x,y)$ represents the neighborhood, J^c stands for each color channel, and J^{dark} represents the dark channel map. Apart from the sky regions, the intensity of J^{dark} is low and close to 0. Combining DCP and ASM, the transmission can be computed by

$$t(x,y) = 1 - \omega \min_{C} \left(\min_{z \in \mu(x,y)} \frac{I(x,y)^C(Z)}{A^C} \right) \tag{8}$$

where ω denotes the fog retention coefficient, which aims to preserve a minimal amount of fog in the distant regions of the original image. In Equation (8), the value of ω is assigned as 0.95 and the atmospheric light is obtained from the region where the darkest 5 percent is. Once these imaging parameters are determined, the clear version can be recovered by Equation (7) from a single foggy image. It should be pointed out that the transmission obtained through this method generally offers high accuracy; however, its high complexity limits its practical application. To address this limitation, He et al. proposed the guided filter (GIF) [11], which improves its computational efficiency. Furthermore, this approach still has some other drawbacks, e.g., it cannot deal with images with sky regions.

2.2.3. Color Attenuation Prior Image Defogging

As is well known, human perception can swiftly discern areas with or without fog, as well as distinguish between near and far distances without relying heavily on additional data. Building upon this observation, Zhu et al. [12] propose a color attenuation prior (CAP) by assuming that haze concentration correlates with the disparity between brightness and saturation. Mathematically, the CAP is expressed by the following:

$$d(x) = \theta_0 + \theta_1 v(x) + \theta_2 s(x) + \varepsilon(x) \tag{9}$$

where θ_0, θ_1, and θ_2 are unknown linear coefficients, and $\varepsilon(x)$ represents a random variable that captures the inherent error of the model. Additionally, $d(x)$, $v(x)$, and $s(x)$ correspond to scene depth, brightness, and saturation, respectively. For simplicity, this approach further assumes a Gaussian distribution for ε with zero mean and variance σ^2 (i.e., $\varepsilon(x) \sim N(0, \sigma^2)$). By leveraging the properties of the Gaussian distribution, Equation (9) can be detailed as

$$d(x) \sim p(d(x)|x, \theta_0, \theta_1, \theta_2, \sigma^2) = N(\theta_0 + \theta_1 v + \theta_2 s, \sigma^2) \tag{10}$$

To learn the coefficients θ_0, θ_1, and θ_2 accurately, Zhu et al. further create a joint conditional concentration in terms of Equation (10), i.e.,

$$L = p(d(x_1), \dots, d(x_n)|x_1, \dots, x_n, \theta_0, \theta_1, \theta_2, \sigma^2) \tag{11}$$

where n is the total number of pixels within the training hazy images, $d(x_n$ is the depth of the nth scene point, and L is the likelihood. The main advantage of this method is its highly effective performance. However, it may be invalid when processing certain cases, e.g., scenes with strong lighting or complex colors.

2.2.4. Gamma Correction Prior to Image Defogging

As discussed above, most currently available methods fail to accurately search the scene depth that is useful for transmission estimation. To this end, Ju et al. [13] firstly introduces a novel pre-processing technique called gamma correction preprocessing (GCP), i.e.,

$$I_s^c = 1 - (1 - I^c)^\Gamma, \tag{12}$$

where I_s is the virtual result, and I^c is the color channel of the hazy image. Having the input image and the prepossessed result, single-image defogging can be subtly changed to multi-image defogging. Taking ASM as theory, the imaging equation of the input image and the prepossessed result is described by

$$I^c(x, y) = A^c \cdot \rho(x, y) \cdot e^{-\beta \cdot d(x,y)} + A^c \cdot (1 - e^{-\beta \cdot d(x,y)}) \tag{13}$$

$$I_s^c(x, y) = A_s^c \cdot \rho(x, y) \cdot e^{-\beta_s \cdot d(x,y)} + A_s^c \cdot (1 - e^{-\beta_s \cdot d(x,y)}) \tag{14}$$

where A^c s is the atmospheric light of the virtual results I^c. By solving this equation, the scene depth can be computed by

$$d = \frac{-\ln \frac{\max(A^c - I^c, \epsilon_1)}{\max(A_s^c - I_s^c, \epsilon_2)} - \ln \frac{A_s^c}{A^c}}{\beta - \beta_s} \tag{15}$$

where ϵ_1 and ϵ_2 are very small positive constants, ϵ_1 is introduced to avoid the numerator from exceeding the function definition field, and ϵ_2 is introduced to make sure the denominator is not zero. For simplicity, Ju et al. further assume that the weather conditions do not change spatially, which leads to

$$d = \frac{1}{\beta - \beta_s} \cdot d_0 \propto d_0 = -\ln \frac{\max(A^c - I^c, \epsilon_1)}{\max(A_s^c - I_s^c, \epsilon_2)} - \ln \frac{A_s^c}{A^c} \tag{16}$$

By substituting Equations (5) and (16) into ASM, the dazed result can be expressed as:

$$\rho^c = \frac{I^c - A^c}{A^c \cdot e^{-\frac{\beta}{\beta - \beta_s} \cdot d_0}} + 1 \tag{17}$$

where the range of ρ^c is set to $0 \leq \rho^c \leq 1$ in order to prevent pixel overflow. Consequently, the modified expression for reconstructing the scene contents can be formulated as follows:

$$\begin{aligned} \rho^c &= \text{dehaze}(\theta, I^c, A^c, d_0) \\ &= \min((\max(\frac{I^c - A^c}{A^c \cdot e^{-\theta \cdot d_0}}) + 1), 0), 1) \end{aligned} \tag{18}$$

where $\text{dehaze}(\cdot)$ is the abbreviation of the albedo restoring function. Note that $\text{dehaze}(\cdot)$ is a function of four parameters, where I^c is the input, A^c can be easily calculated according to Equation (18), d_0 is the depth ratio obtained in the previous subsection, and $\theta = \beta/(\beta - \beta_s)$ is the only unknown parameter. To estimate the value of θ with low complexity but high accuracy, a globally optimized function is designed as:

$$\theta = \text{argmin}\{\sum_c f(\text{dehaze}(\theta, (I^c) \downarrow^n, A^c, (d_0) \downarrow^n))\} \tag{19}$$

where $f(\cdot)$ represents a vision indicator designed via single or multiple images prior, and $\downarrow^n$ is a down-sampling operator with coefficient n. With this estimated θ, the clear version can be directly recovered according to Equation (18). Unlike the other defogging methods employing pixel-wise, patch-wise, scene-wise, non-local-wise, and learning-wise strategies, this technique makes use of a global-wise strategy to achieve a high-quality image defogging. Nevertheless, because of the fact that it assumes weather conditions do not change spatially, thus making it fails to deal with the images with non-homogeneous fog.

2.2.5. Physical-Model-Based Defogging Using Other Prior Knowledge

Tan [15] assumes a fixed atmospheric light value in the local region and employs a Markov model framework to maximize local contrast for processing foggy images. This is achieved by developing a cost function and estimating the optimal atmospheric light using graph segmentation knowledge. The algorithm effectively enhances image contrast and improves visibility; however, it may lead to color over-saturation post-fog removal and introduce halo effects in certain interface areas.

Fattal [16] assumes that the local region's reflectance remains constant, while the object's surface chromaticity demonstrates local statistical intercorrelation with media propagation. Nevertheless, accurate estimation can pose challenges in cases where relevant components lack noticeable changes or when color information is limited.

Tarel et al. [17] proposed a fast fog removal algorithm, which estimates the dissipation function by analyzing the distortion caused by the median filter. Regrettably, inappropriate parameter configurations during the application of the median filter estimation method can introduce halo artifacts.

Ju et al. [18] explores a region line prior (RLP), i.e., when the image is divided into n regions, the brightness corresponding to the blurred image and the fog-free image in each region is positively correlated with the scene depth. Then, combining RLP and ASM, they further proposed the defogging algorithm. To solve the dim effect and better simulate the outdoor hazy scene, Ju et al. [43] also developed a simple yet effective image enhancement technology based on the grayscale world hypothesis and an enhanced ASM.

Berman et al. [19] proposed the non-local image defogging algorithm based on the assumption that the color of the haze-free image can be well estimated from hundreds of strict colors. Since this algorithm is based on pixels rather than patches, it can have a high efficient and exhibit a high-quality enhancement effect.

Oakley et al. [20] postulated the availability of scene depth information and utilized Gaussian functions to restore scene contrast by predicting the optical path. Importantly, their approach did not necessitate any weather-related predictions. However, the imple-

mentation conditions were demanding, requiring specific hardware devices for obtaining depth-of-field (DOF).

Kopf et al. [21] employed a combination of hardware and software devices to acquire auxiliary information, thereby facilitating the collection of depth-of-field (DOF) and texture data. Despite their development of a novel system, it still fails to address the limitations associated with the requirement for specific equipment to obtain DOF.

3. Deep Learning Defogging

Deep learning technologies have made significant advancements in the field of image defogging. Depending on the deep networks or training strategies used during learning stage, in this section, we will introduce the deep learning fog removal algorithm in two parts, i.e., the architecture of a deep network and training strategy used for image defogging.

3.1. Architecture of Deep Network Used for Image Defogging

Currently available deep learning fog removal methods select different deep networks (e.g., Convolutional Neural Network (CNN), Generative Adversarial Network (GAN), Residual Network (ResNet), Attention Mechanism Network, and autoencoder) as their backbone to implement fog removal, depending on their concerns, as shown in Figure 2.
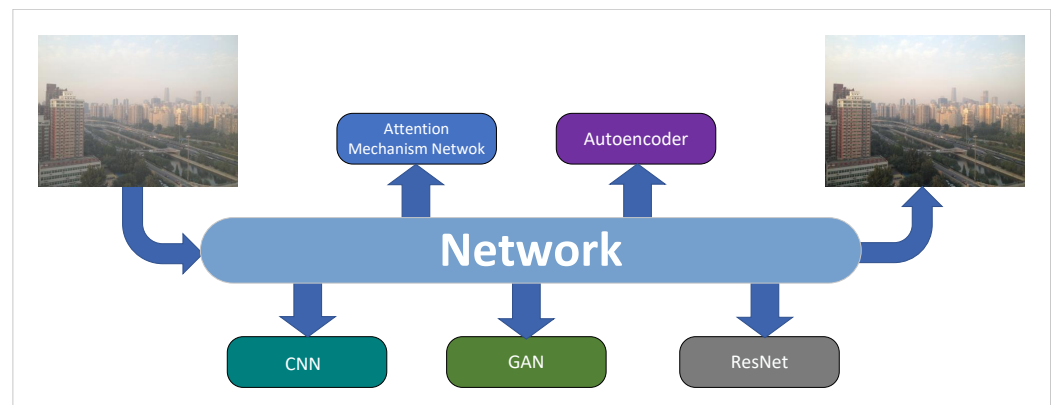

Figure 2. The network architecture diagram of deep learning defogging algorithms.

3.1.1. CNN-Based Defogging Methods

CNN is well-suited for image defogging, based on its capability at capturing local features by using filters and feature maps. Such abilities enable the network to better learn edges, textures, and other low-level visual elements. Pooling layers then reduce the spatial dimensions of the images and realize computational efficiency. Finally, the fully connected layers integrate the high-level features learned by the previous layers, enabling the network to make complex decisions and classifications. One of the most typical networks using CNN, DehazeNet [44], is the first deep learning-based method for image defogging. It estimates the medium transmission map of a hazy image through an end-to-end trainable system. The system employs a CNN architecture, with layers designed to embody established assumptions and priors. The core contributions of DehazeNet are as follows:

- End-to-end learning: The mapping relations between hazy patches and their medium transmissions can be directly learned and estimated by DehazeNet. Such an end-to-end learning method brings strong flexibility to the method.
- Novel nonlinear activation function: DehazeNet proposes a new nonlinear activation function called the Bilateral Rectified Linear Unit (BReLU) to improve the quality of restored haze-free images. BReLU reduces the search space and enhances convergence through bilateral restraint.

- Connection with traditional dehazing methods: DehazeNet integrates its components with the assumptions/prior knowledge used in existing defogging methods, which makes it better performing compared to the traditional methods.

DehazeNet's architecture includes multiple convolutional and pooling layers, as well as Maxout and BReLU activation functions. The medium transmission is estimated through four steps:

1. Feature extraction: maximize unit extraction features and generate almost all haze-relevant features.
2. Multi-scale mapping: to achieve scale invariance, multi-scale features are extracted through parallel convolutional operations.
3. Local extremum: local extremum operations are employed to achieve spatial invariance, aligning with the assumption of locally constant medium transmission.
4. Nonlinear regression: nonlinear regression is conducted using the BReLU activation function to restrict transmission values within a reasonable range, which can effectively decrease noise issues.

CNN has become a foundation in image defogging due to its powerful feature extraction capabilities and end-to-end ability. However, the demand for efficiency and effectiveness continues to drive the development of newer models. One such recent development is the Light-DehazeNet [45], which stands out for its ability to deliver high-quality defogging results with improved computational efficiency. This algorithm involves the following advancements:

- Efficient design: Light-DehazeNet is designed to be lightweight, making it much more practical for application.
- Advanced feature learning: it adopts complex yet efficient convolutional layers that can capture detailed features, even in low-visibility conditions.
- Seamless integration: the model's structure allows for seamless integration with existing CNN-based pipelines, which can ensure compatibility with traditional methods and offer enhanced performance at the same time.
- Real-time capability: with its streamlined processes, Light-DehazeNet is capable of real-time defogging, which is a significant advantage for applications requiring instant visual clarity.

While previous studies have made significant advancements in the field of image defogging through its lightweight architecture and end-to-end training strategy, further research is still required. This includes integrating the atmospheric scattering model directly into deeper neural networks and exploring more complex network structures to enhance defogging performance.

3.1.2. GAN-Based Defogging Methods

GAN [46] has revolutionized deep learning by offering a unique approach to generative modeling. Introduced by Ian Goodfellow and colleagues in 2014, GAN consists of two neural networks, a generator and a discriminator. The goal of the generator is to produce synthetic data that mimics real data, whereas the discriminator's role is to distinguish between genuine and generated data. Hence, a network trained in this way can output almost real images. The intricate dynamics within GANs and their ability to capture complex data distributions make them a compelling subject of study in the application of image defogging.

For example, Cycle-Dehaze is an end-to-end network designed specifically to address the single image defogging without the requirement for paired hazy and clear images for training. Additionally, Cycle-Dehaze does not need to estimate atmospheric scattering model parameters.

To generate visually superior haze-free images, Cycle-Dehaze improves upon the CycleGAN formulation by incorporating cycle consistency and perceptual losses. A perceptual loss based on the VGG16 feature extractor is introduced to preserve the original

image structure by comparing images in the feature space rather than the pixel space. To decrease image quality deterioration during the defogging process, the Laplacian algorithm is used for up-sampling. Building upon the foundation laid by earlier GAN-based dehazing methods, like the Cycle-Dehaze mentioned above, the ADE-CycleGAN [47] introduces a significant leap forward in the preservation of image details. This approach addresses the limitations of previous techniques where fine details were often lost during training. Its contributions are as follows:

- Enhanced detail preservation: by integrating a multi-head attention mechanism within a CycleGAN framework, ADE-CycleGAN effectively captures and retains the intricate details of hazy images.
- Improved performance metrics: the novel structure demonstrates a substantial improvement over traditional CycleGAN in quantitative assessments, reflecting a higher accuracy in the reconstructed images.
- Adversarial training: ADE-CycleGAN employs adversarial training to ensure that the dehazed images are not only detail-rich, but also visually realistic.
- Cyclic consistency: the framework also incorporates cyclic consistency loss, ensuring that the dehazed image remains true to the original, thus avoiding artifacts that older GAN models might introduce.

The introduction of ADE-CycleGAN signifies a refinement in GAN-based dehazing techniques, offering a more nuanced and effective approach to recovering clear visuals from hazy scenes.

3.1.3. ResNet-Based Defogging Methods

Residual Networks (ResNets) [48], as a type of deep neural network structure, have demonstrated outstanding performance in image-defogging tasks. Figure 3 illustrates the structure diagram of a ResNet. Its main function is to address the common issue of vanishing gradients during the training process of deep neural networks by introducing residual blocks, making the network easier to train and enabling it to learn image features more deeply.
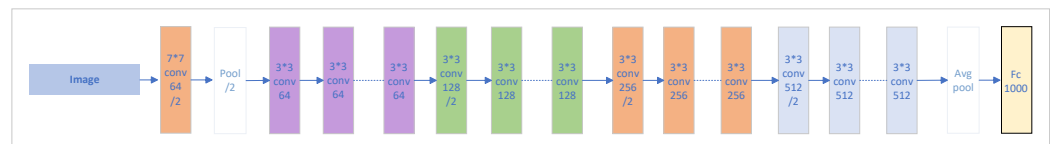

Figure 3. A visualization to illustrate ResNet.

To apply ResNet in image defogging, researchers have designed various specialized defogging networks that incorporate residual blocks to enhance defogging performance. This network structure excels not only in removing atmospheric scattering effects, but also in capturing deep image features, contributing to the preservation of more details, and improving the visual quality of the images. For instance, the most representative of this class is AOD-Net [49] proposed by Boyi Li et al. at the 2017 IEEE International Conference on Computer Vision (ICCV). It employs a Convolutional Neural Network (CNN) approach, particularly adapting the design philosophy of ResNet. The core advantage of AOD-Net lies in its end-to-end learning. AOD-Net generates clear images directly from hazy images through a lightweight CNN, eliminating the need for additional post-processing steps. This end-to-end learning approach reduces mistakes that might be introduced by intermediate stages. As a result, the defogging performance is significantly enhanced. Moreover, AOD-Net has high processing speed, with the ability to process a 480 × 640 pixel image in approximately 0.026 s, making it suitable for real-time applications.

3.1.4. Attention Mechanism Network-Based Defogging Methods

An Attention Mechanism aims to enable the network to focus more on crucial regions of the image, allowing for more effective extraction and utilization of key information. Its core idea is to simulate the human visual system's attention mechanism, enabling the network to concentrate on significant parts of the image during processing. Such a mechanism allocates varying weights to different layers or channels in the network. This personalized attention capability makes the network more flexible, and allows it to better adapt to image-defogging tasks with varying complexity and feature distributions. Therefore, the application of Attention Mechanism Networks brings new advancements to the field of image defogging, offering innovative solutions for addressing atmospheric scattering issues in real-world scenarios.

One of the most representative methods of using an attention mechanism network, GridDehazeNet [50], an attention-based multi-scale Convolutional Neural Network, has demonstrated its effectiveness in single-image defogging tasks. The network consists of three main modules: a pre-processing module, a backbone network module, and a post-processing module.

- **Pre-processing module**: This fully trainable module generates learned inputs that are more diverse and relevant compared to those derived from hand-selected pre-processing methods. This learned pre-processing approach better highlights different aspects of the image, providing richer information for subsequent defogging operations.
- **Backbone network module**: This is the core of GridDehazeNet. It implements a novel attention-based multi-scale estimation method on a grid network. The structure allows for dense connections between different scales, effectively relieving the bottleneck issues commonly encountered in conventional multi-scale approaches. Each scale consists of a series of residual dense blocks (RDBs) connected by upsampling or down-sampling blocks, enabling flexible aggregation and fusion of features across scales.
- **Post-processing module**: To address the issue that features from different scales may have varying weights, GridDehazeNet incorporates a channel-wise attention mechanism. This mechanism allows the network to generate trainable weights for feature fusion, thereby flexibly adjusting the contributions from different scales and further improving defogging quality.

Grid-DehazeNet showcases the potential of attention mechanisms in image-defogging tasks, particularly in handling multi-scale features and enhancing defogging quality. This network design provides a new perspective for the field of image defogging and may inspire future research on applying attention mechanisms in other image restoration tasks.

3.1.5. Autoencoder-Based Defogging Methods

Autoencoder [51], a crucial neural network structure in deep learning, has demonstrated remarkable performance in image defogging tasks. The design of autoencoder aims to learn a concise representation of data through unsupervised learning, enabling efficient information reconstruction between input and output. The core structure of an autoencoder comprises an encoder and a decoder. The encoder maps the input image to a low-dimensional representation space, while the decoder reconstructs this low-dimensional representation into a reconstructed image. This structure enables the autoencoder to capture essential features in image defogging tasks, reducing noise and blur effects during the reconstruction process, and thereby enhancing image clarity and quality.

In the application of image defogging, the network structure of autoencoder is often adjusted by modifying the layers and nodes of the encoder and decoder to accommodate different tasks and datasets. For example, an autoencoder-based Contrastive-Regularized Network (AECR-Net) is a defogging network that integrates concepts from autoencoders, single image defogging, and contrastive regularization. The network employs a compact framework that includes an autoencoder for feature extraction and reconstruction, which is beneficial for memory storage and performance balance.

The key innovations of AECR-Net is the introduction of a novel Contrastive Regularization (CR) technique. This technique is grounded in contrastive learning, which uses both the hazy and clear images as negative and positive samples, respectively. The CR aims to improve the network's ability to distinguish between the features of hazy and clear images, thus enhancing the defogging outcome. The network architecture of AECR-Net is compact, making it suitable for applications where computational resources are limited. Despite its compactness, AECR-Net can produce high-quality dehazed images, making it a promising approach in the field of image processing for defogging tasks.

While AECR-Net and related research during the same period have laid the groundwork for feature learning in image defogging, the latest developments have pushed the boundaries further. One such breakthrough is the RIDCP [52], which integrates the concept of high-quality codebook priors into the autoencoder framework. Unlike earlier methods that relied on hand-crafted features and heuristic models, RIDCP leverages the power of VQGAN to encapsulate robust priors directly from a large-scale dataset of pristine images. Its innovations are as follows:

- **High-quality codebook priors**: By pre-training on a diverse set of high-quality images, RIDCP creates a discrete codebook that serves as a rich source of prior knowledge for the defogging task. This approach contrasts with older methods that often struggled with limited training data and less representative features.
- **Phenomenological data synthesis**: addressing the synthetic-real gap, RIDCP introduces a synthesis pipeline that mimics real-world degradation more closely than traditional models, enhancing the generalizability of the learned features.
- **Controllable priors matching**: a novel aspect of RIDCP is its controllable matching mechanism, which refines the alignment between hazy image features and the learned priors, leading to improved defogging outcomes compared to the more rigid matching strategies of older autoencoders.

The introduction of RIDCP signifies a shift towards more sophisticated autoencoder models that can better handle the intricacies of real-world defogging. It builds upon the foundational work of previous autoencoder techniques, taking a significant leap forward in terms of performance and adaptability.

3.2. Training Strategy Used for Image Defogging

Currently available deep learning fog removal methods utilize a supervised mode, unsupervised mode, and semi-supervised mode to train their constructed network. Figure 4 showcases a diagram of how different training modes are embedded in the network. It is easily concluded from this figure that these available networks either employ paired datasets or unpaired datasets to fit the pre-set deep models.
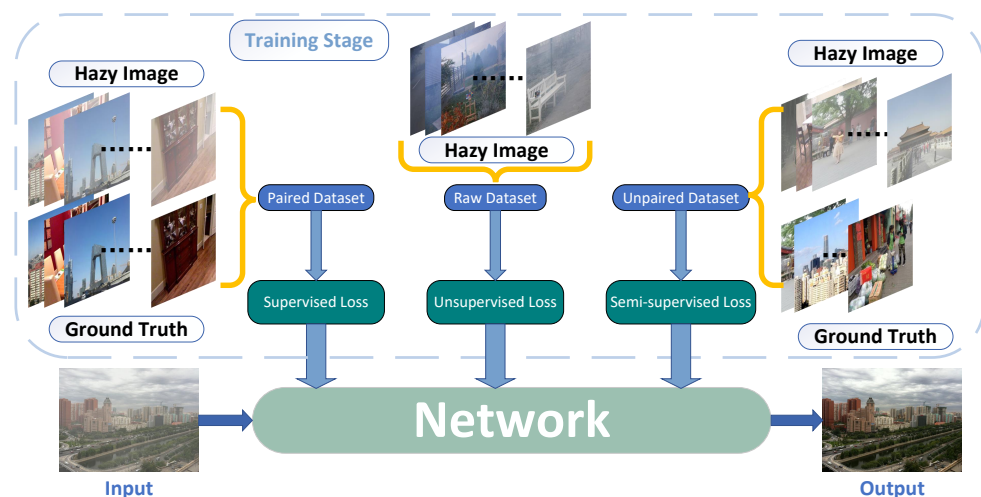

Figure 4. Example description of how different training modes are embedded in the network.

3.2.1. Image Defogging Based on Supervised Learning Mode

Supervised defogging algorithms aim to enhance visibility of hazy images by leveraging the labeled training data. The basic steps involve collecting a dataset of hazy and corresponding clear images, extracting features by each layer in deep network, training the designed model using supervised learning loss, and finally applying the trained model to achieve image fog removal. Through this process, the model learns to estimate and remove haze by understanding the relationship between hazy and clear image pairs. In general, such supervised methods can achieve better performance compared to unsupervised approaches that do not require image pairs, particularly in challenging conditions with varying haze levels and complex scenes.

3.2.2. Image Defogging Based on Unsupervised Learning Mode

Although the aforementioned supervised image-defogging methods have demonstrated outstanding results, their performance heavily relies on paired datasets during training. To address the challenge of dataset shortage, the unsupervised learning image-defogging methods have garnered significant attention. The unsupervised method can effectively reduce the cost of data acquisition, which makes it more suitable for real-world defogging tasks. To the best of our knowledge, the unsupervised image defogging can be roughly divided into unsupervised learning approach based on prior knowledge and unsupervised learning approach based on generative adversarial network.

Unsupervised learning based on prior knowledge: The unsupervised learning approach based on prior knowledge involves combining, transforming, and deriving feature information through domain knowledge related to physics and statistics. Such knowledge not only helps to improve generalization capabilities, but also reduces over-fitting during training. For example, Golts et al. [22] proposed a method combining dark channel prior (DCP) and unsupervised fog combination. They tune the network's parameters directly minimizing DCP on foggy images, ensuring color fidelity and generalization ability. In ref. [23], Sham et al. presented a prior-based adversarial training approach to ensure consistent performance on synthetic and real datasets. This method efficiently leverages both low-frequency and high-frequency components of an image to safeguard critical color and structural details in the recovered image, resulting in improved generalization capabilities and reliable performance in practical scenarios. Golts et al. [24] proposed an simple but effective approach that trains deep neural networks (DNNs) through the minimization of energy functions rather than generic loss functions.

Unsupervised learning based on GAN: In addition to the previously mentioned unsupervised strategies, there is another type of network worth paying attention to, namely GAN. The model is trained by two networks, the generator and the discriminator. The model was trained through the constant confrontation between the generator and the discriminator. In unsupervised single image dehazing with generative adversarial network [28], they proposed a network consisting of a generator, a global test discriminator, and a local context discriminator, combined with an attention mechanism based on a dark channel prior to dealing with unevenly distributed haze. In AAGAN [29], an attention-to-attention generative adversarial network (AAGAN) is presented for single image defogging. This network incorporates a dense channel attention model within the encoder and a multi-scale spatial attention model within the decoder, aiming to enhance the defogging performance. This attention-to-attention model can better extract the global features and achieve better defogging performances. The researchers further introduce spectral normalization of all convolutional layers to stabilize the training process. Most deep learning methods for image defogging tend to treat hazy images regardless of their varying degrees of haziness as belonging to the same image domain, overlooking the significant domain differences that arise from different haze densities. In Discrete Haze Level Dehazing Network [30], the discrete haze level defogging network (DHLDehaze) was proposed, aiming to process multiple haze level domains. Leveraging the adversarial training process involving the source and target domains enables the network to transform the haze level of an image while preserving the integrity of its scene content.

3.2.3. Semi-Supervised Image Dehazing Methods

Due to the lack of paired data, researchers also develop several semi-supervised image-defogging methods. The mainstream framework of semi-supervised image-defogging methods mainly contains two branches: the supervised training branch and the unsupervised training branch. The main difference between the supervised branch and the unsupervised branch is mainly reflected in the setting of the loss function. In Semi-Supervised Image Dehazing [37], an algorithm employing a deep CNN is presented, comprising a supervised learning branch and an unsupervised learning branch. The semi-supervised approach performs better than the unsupervised approach in completely eliminating the haze, but it is prone to color shifts and stains. To address the issue, a channel-spatial self-attention (CSSA) mechanism composed of three parts: channel attention, spatial attention, and self-attention were introduced to the network, called SAD-Net, proposed in Semi-Supervised Domain Alignment Learning for Single Image Dehazing [38]. The CSSA method flexibly assigns weights to features that enable the network to better extract important information, thereby enhancing the network's defogging performance. To enhance generalization in real-world scenarios, researchers [38] introduced a domain alignment module and a haze-aware attention module into their network architecture. These modules help the network to narrow the distribution distance between synthetic data and realistic hazy images in a latent feature space, and adaptively respond to different hazy areas.

4. Experiment

In this section, to intuitively illustrate the advantages and disadvantages of different types of algorithms, their recovery performances are evaluated from qualitative and quantitative perspectives. First, we conducted a comparison of classic traditional methods (HE, AHE, CLAHE, Homomorphic Filtering [53], SSR, MSR, and Laplace [54]) to reveal the shortcomings of simply increasing contrast (here, we remark that, despite the fact that these traditional methods comes from older literature sources, they represent the early evolution of image enhancement). Then, the results restored by physical-model-based methods (CAP [12], DCP [10], IDE [43], IDRLP [18], NLP [19], and TERAL [17]) were evaluated and compared. Subsequently, we quantitatively and qualitatively compared the results obtained by state-of-the-art deep learning-based technologies, including deep models employing different networks (AOD, Cycledehaze, DehazeNet, GridDehazeNet) and deep models using different training strategies (C2P-Net, FFA, Taylor, UHD, Vison, DE, USID, SLA, SDA-GAN), on various challenging hazy images. For fairness, the codes of selected available techniques for comparison were downloaded from authors' homepage, and the parameters used in these techniques were optimized according to the corresponding references. Note that all of the experiments were implemented on a PC with an Intel(R) Core(Tm) i5-4210U CPU @ 1.70 GHz, 8.00 GB RAM, and NVIDIA 3090 Ti GPU (for comparative algorithms, their more detailed configurations as shown in Table 1), and the hazy images used in the experiments were collected from publicly available datasets (I-haze [55], O-haze [56], and SOTS [57]).

Table 1. The list of the the classical and state-of-the-art methods used in our experiments.

Methods Selected for Experiments
Non-deep learning defogging:
Traditional Methods: **HE**: ICCCNT, 2013 (CPU)   **AHE**: APCC, 2019 (CPU)   **CLAHE**: IEEE TIP, 2013 (CPU)   **Homomorphic Filtering**: Pattern Recognition Letters, 2011 (CPU)   **SSR**: ICDIIME, 2023 (CPU)   **MSR**: CAC, 2020 (CPU)   **Laplace**: McGraw-Hill, 1965 (CPU)

Table 1. *Cont.*

Methods Selected for Experiments
Physical-Model-Based Methods: **CAP**: IEEE TIP, 2015 (CPU)   **DCP**: IEEE CVPR, 2009 (CPU)   **IDE**: IEEE TIP, 2021 (CPU)   **IDRLP**: IEEE TIP, 2021 (CPU)   **NLP**: IEEE CVPR, 2016 (CPU)   **TERAL**: IEEE ICCV, 2009 (CPU)
Deep Learning-Based Methods:
Deep Learning-Based Methods of Network Architecture: **AOD**: IEEE ICCV, 2017 (CPU) (ResNet)   **Cycledehaze**: IEEE CVPR, 2018 (CPU) (GAN)   **DehazeNet**: IEEE TIP, 2016 (CPU) (CNN)   **GridDehazeNet**: IEEE ICCV, 2019 (CPU) (Attention)
Deep Learning-Based Methods of Training Mode: **C2P-Net**: IEEE CVPR, 2023 (GPU) (Supervised)   **FFA**: AAAI, 2020 (GPU) (Supervised)   **Taylor**: IEEE ICCV 2023 (GPU) (Supervised)   **UHD**: IEEE CVPR 2021 (GPU) (Supervised)   **Vison**: IEEE TIP 2023 (GPU) (Supervised)   **DE**: IEEE J-STSP 2021 (CPU) (Unsupervised)   **USID**: IEEE TIP 2020 (CPU) (Unsupervised)   **SLA**: IJCAI 2022 (CPU) (Semi-supervised Learning)   **SDA-GAN**: IEEE T CYBERNETICS 2023 (CPU) (Semi-Supervised)

4.1. Performance Description of Non-Deep Learning Defogging

As discussed above, non-deep learning defogging mainly includes two types: traditional image defogging and physical-model-based defogging. To investigate the advantages and disadvantages of these two types of algorithms, we conducted a lot of experiments based on O-haze, D-haze [42], I-haze, and DN-haze [58] datasets. Here, we remark that, the reason of using these datasets is because they contain different real-world scenes with different haze thickness distributions, which can better check the performance of different algorithms.

4.1.1. Limitations of Traditional Image Defogging

Qualitative comparison: In this subsection, seven representative techniques, i.e., HE, AHE, CLAHE, Homomorphic Filtering, SSR, MSR, and Laplace, were selected to check their performance on a variety of challenging hazy images. The comparison results are illustrated in Figure 5. As seen in Figure 5, HE is capable of dealing with most scenes. However, it may lead to darker performances and suffer from severe artifacts. AHE may suffer from the color cast issue, and CLAHE still leaves a significant amount of residual haze. The Homomorphic Filtering method causes severe color distortion on given examples, and yields disastrous results when processing high-brightness regions. SSR and MSR exhibits over-enhancement of the sky regions and color distortions in misty scenes. Although Laplace effectively enhances edges to a certain extent, it introduces noise and blurring.

Quantitative comparison: To reach a more comprehensive evaluation, Peak Signal-to-Noise Ratio (PSNR) [59] and Structural Similarity (SSIM) [60] calculated on several representative algorithms (HE, AHE, CLAHE, Homomorphic Filtering, SSR, MSR, and Laplace approaches) based on the SOTS dataset are shown in Table 2. As seen from this table, all of the selected traditional image enhancement algorithms can have a fast processing speed. However, their scores of PSNR and SSIM still present several low values, which reveals that these methods lack the ability to remove the haze cover in an images. Taking CLAHE as an example, for a few examples, this method can eliminate the effect caused by haze

thanks to its local contrast enhancement capability, while its enhanced results appear to blur and color cast.

Figure 5. Comparison figure of image enhancement-based traditional defog algorithms on synthetic images.

Table 2. Comparison table of image enhancement-based traditional defog algorithms.

Classification	Metrics	HE	AHE	CLAHE	HF	SSR	MSR	LAP
indoor	PNSR	12.88	15.20	**15.69**	12.71	7.67	8.34	12.01
	SSIM	0.7182	0.7603	**0.7765**	0.7140	0.5554	0.5781	0.6657
outdoor	PNSR	16.09	18.23	**19.88**	17.11	10.63	11.63	15.21
	SSIM	0.8201	0.8300	**0.8684**	0.8668	0.7155	0.7438	0.6990
	Runtime (s)	0.0296	0.0303	0.0522	0.0761	0.0921	0.5929	**0.0068**

1. HF represents Homomorphic Filtering. 2. LAP represents Laplace operator. 3. The bold data is optimal data.

4.1.2. Limitations of Physical-Model-Based Defogging

Qualitative comparison: Figure 6 shows the results dehazed by six representative techniques, including CAP, DCP, IDE, IDRLP, NLP, and TERAL. Note that these images used for comparison were also picked from the above dataset. As shown in Figure 6, the DCP-based approach shows its advantages on different datasets. However, the DCP algorithm can lead to over-saturation and color distortion. For other mainstream algorithms, CAP, IDE, and IDRLP all overmagnify the details of image content, and thus produce some undesirable artifacts in the dehazed results.

Quantitative comparison: To obtain a more reliable conclusion, the calculated PSNR and SSIM for CAP, DCP, IDE, IDRLP, NLP, and TERAL algorithms are summarized in Table 3. By comparison, it can be found that IDRLP has the best performance in the test results of the four data sets. In particular, its PSNR and SSIM indicators are excellent, and the running time is very short. However, this does not mean that it can serve as an excellent candidate for the image fog removal task. This is because the prior knowledge employed by these physical-model-based defogging methods may fail in some cases, thereby resulting in some negative effects in enhanced versions.
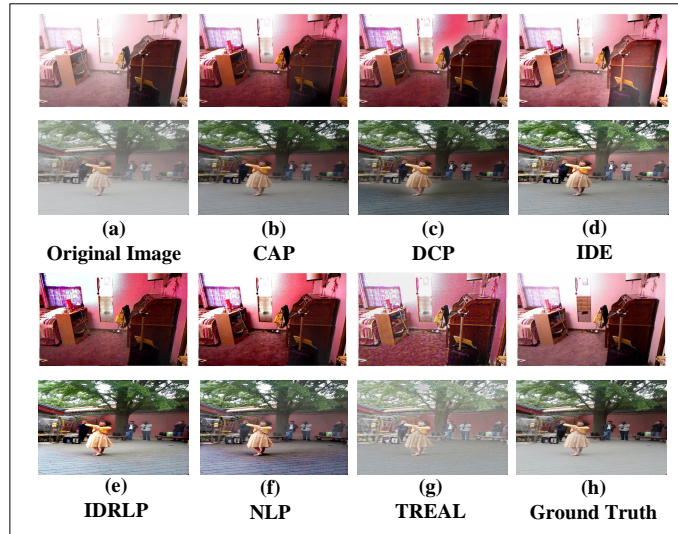
Figure 6. Comparison figure of physical-model-based traditional defog algorithm on synthetic images.

Table 3. Comparison table of physical-model-based defogging algorithms.

Dataset	Metrics	CAP [12]	DCP [10]	IDE [43]	IDRLP [18]	NLP [19]	TREAL [17]
I-haze	Runtime (s)	9.88	2.70	2.544	3.901	20.672	**2.505**
	SSIM	0.779	0.617	0.72	**0.789**	0.706	0.595
	PSNR	13.405	11.303	15.917	**17.355**	13.580	10.403
	FADE	1.042	0.625	0.525	1.083	0.630	**0.279**
O-haze	Runtime (s)	10.299	2.438	2.170	2.352	27.709	**1.491**
	SSIM	0.715	**0.720**	0.632	0.699	0.621	0.654
	PSNR	15.29	14.906	14.299	**16.949**	12.686	12.563
	FADE	0.754	0.379	0.460	0.698	0.527	**0.135**
SOTS-In	Runtime (s)	0.620	0.065	0.857	0.371	3.440	**0.053**
	SSIM	0.799	0.779	0.798	**0.901**	0.769	0.766
	PSNR	18.936	19.525	17.324	**20.100**	17.282	17.006
	FADE	0.590	0.497	0.645	0.334	0.474	**0.241**
SOTS-Out	Runtime (s)	0.584	0.065	0.877	0.220	3.115	**0.037**
	SSIM	**0.850**	0.810	0.820	0.847	0.770	0.741
	PSNR	18.117	17.215	18.782	**19.245**	17.975	16.987
	FADE	0.541	0.448	0.613	0.418	0.481	**0.229**

1.These methods are processed on Matlab(CPU) 2.The bold data is optimal data.

4.2. Performance Description of Deep Learning Defogging

4.2.1. Performance Analysis of Network Architecture

Qualitative comparison: In this subsection, we further checked the performance among different image defogging methods, including DehazeNet (using CNN architecture), GridDehazeNet (using Attention architecture), AOD (using ResNet architecture), and Cycledehaze (using GAN architecture), on various challenging synthetic images. The corresponding experimental results of the dehazing models using different architectures are shown in Figure 7. It is easily noted from this figure that, regardless of the network architecture used, it seems that deep learning defogging would have a better processing performance than non-deep learning defogging techniques. The key to achieve this success can be attributed to the strong fitting ability of deep models. However, these architectures still have their drawbacks, e.g., CNN can not work well on dark regions and GAN fails to deal with sky parts.
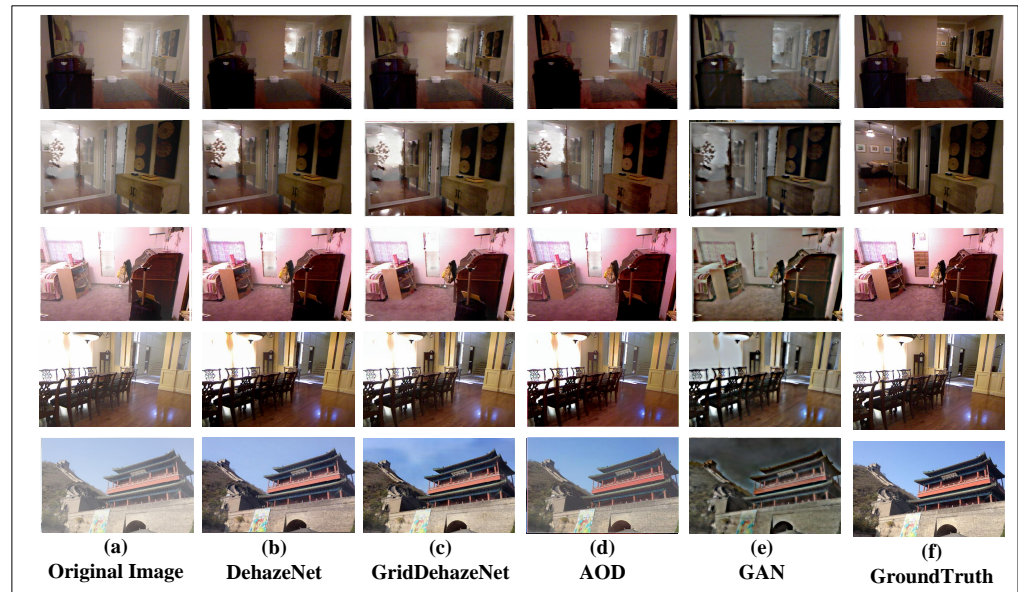
Figure 7. Comparison figure of different deep learning models on synthetic images.

Quantitative comparison: To obtain a more reliable conclusion, the calculated PSNR and SSIM for DehazeNet, GridDehazeNet, AOD, and Cycledehaze are summarized in Table 4. Note that these metric scores were averaged over all the results from the used datasets. Upon comparison, it is obvious that DehazeNet using CNN architecture generally attains the best scores, while CycleDehaze employing GAN only gets the last place. This means the unpaired dataset trained by GAN can not afford the adequate fog features to defogging models. On the other hand, the values in this table are also further evident that deep learning defogging outperforms non-deep learning defogging techniques in terms of visual quality and quantitative scores.

Table 4. Comparison table of the effectiveness of image enhancement-based traditional defog algorithms.

Dataset	Metrics	DehazeNet [44]	GridDehazeNet [50]	AOD [49]	CycleDehaze [46]
I-haze (640 × 480)	Runtime (s)	**0.273**	0.375	0.788	13.911
	SSIM	**0.841**	0.721	0.641	0.702
	PSNR	**19.115**	15.050	14.864	18.181
	FADE	**0.922**	1.089	1.315	1.035
O-haze (640 × 480)	Runtime (s)	1.578	**0.380**	0.788	11.976
	SSIM	**0.729**	0.604	0.641	0.655
	PSNR	15.236	15.570	14.864	**23.903**
	FADE	0.949	**0.714**	1.315	0.898
SOTS-In	Runtime (s)	1.476	**0.265**	0.788	8.069
	SSIM	**0.864**	0.588	0.641	0.718
	PSNR	**19.976**	14.280	14.864	16.405
	FADE	0.588	**0.395**	1.315	0.527
SOTS-Out	Runtime (s)	1.218	**0.187**	0.273	8.329
	SSIM	**0.885**	0.871	0.841	0.767
	PSNR	**22.923**	22.730	19.115	12.446
	FADE	0.722	**0.456**	0.922	0.804

The bold data is optimal data.

4.2.2. Performance Analysis of Training Mode

Qualitative comparison: In the above, we have experimentally shown that the impact of different network architectures on defogging performance. In fact, the training mode is also crucial to image fog removal models. Therefore, nine state-of-the-art methods

(i.e., UHD, C2P-Net, Taylor, Vision, FFA, USID, DE, SLA, and SDA-GAN) were selected to check the impact of different training mode. The corresponding results dehazed by different available techniques are given in Figure 8. As expected, most of the selected methods can exclude the haze cover in an image to some extent. However, they all have their own limitations. For the methods using supervised mode, they are hard to balance the enhancement quality between thickness haze and mist images. For the methods exploiting unsupervised mode, they lack the ability to handle the regions where the brightness is similar to atmospheric light.
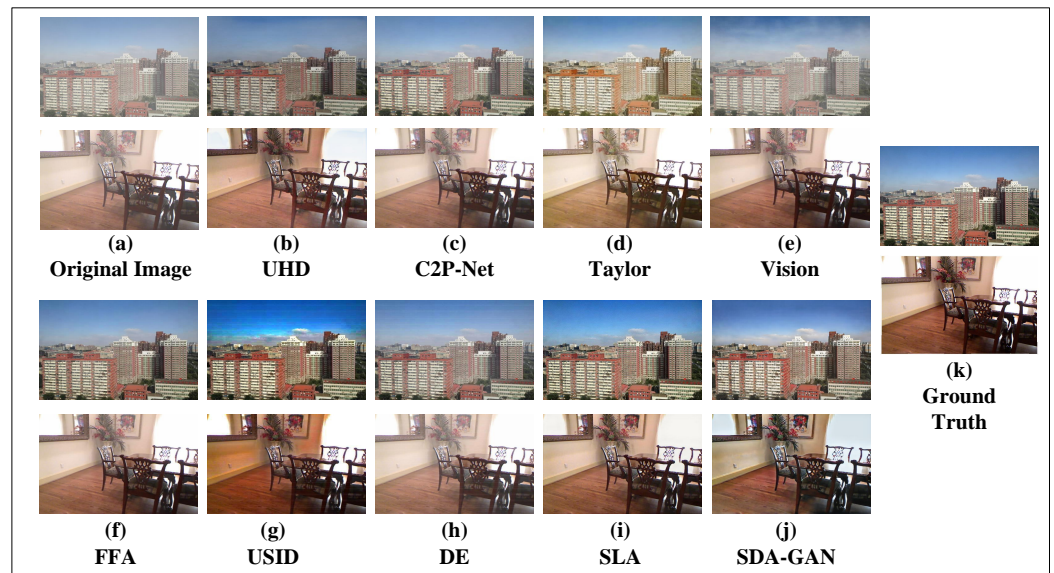

Figure 8. Comparison figure of supervised, unsupervised and semi-supervised methods on synthetic images.

Quantitative comparison: To provide a more comprehensive evaluation, the calculated PSNR and SSIM for the nine algorithms are summarized in Tables 5 and 6. As analyzed from the tables, the defogging networks using the supervised mode have a more robust capability than the ones using the unsupervised mode. However, their computational complexity is significantly higher than that of ones using unsupervised mode and semi-unsupervised mode.

Overall, non-deep learning defogging methods either leverage statistical image properties (traditional image enhancement) or a combined atmospheric scattering model and prior knowledge (physical-model-based defogging) to realize image fog removal. This make them relatively straightforward to implement with low algorithmic complexity and less computational resource consumption. However, because of the fact that the limitations of statistical image properties and prior knowledge, they always fail to deal with complex scenes, especially for the images with uneven fog. For deep learning defogging, different network architecture used in defogging algorithms may exhibit different defogging effects, e.g., CNN is able to effectively extract local features, and transformer can excavate the global features to enhance a single foggy image. Moreover, currently available defogging networks generally make use of supervised, unsupervised, and semi-supervised modes to train the created network. According to the experiment results, the networks using supervised mode can have a reliable performance on synthetic dataset, while they may invalid in real-world scenarios. On the contrary, the models employing unsupervised and semi-supervised modes work well on scenes collected from the real world, yet they fail to process the synthetic images.

Table 5. Comparison table of supervised wise defog algorithms.

Dataset	Metrics	UHD [32]	C2P-Net [31]	Taylor [61]	Vision [36]	FFA [34]
I-haze (640 × 480)	Runtime (s)	0.347	1.016	1.089	**0.111**	0.641
	SSIM	**0.813**	0.675	0.628	0.806	0.642
	PSNR	**20.186**	14.853	14.422	18.008	17.205
	FADE	1.130	0.735	**0.440**	1.306	3.054
O-haze (640 × 480)	Runtime (s)	0.325	0.971	1.091	**0.112**	0.623
	SSIM	0.662	0.600	**0.772**	0.660	0.621
	PSNR	15.322	14.720	**16.565**	12.901	12.899
	FADE	0.784	**0.382**	0.615	0.933	0.768
SOTS-In	Runtime (s)	0.520	0.785	0.841	**0.110**	0.380
	SSIM	0.779	0.508	0.870	0.851	**0.968**
	PSNR	18.320	14.367	26.295	21.924	**31.922**
	FADE	0.733	0.393	0.436	0.451	**0.400**
SOTS-Out	Runtime (s)	0.476	0.716	0.711	**0.116**	0.320
	SSIM	**0.905**	0.822	0.899	0.811	0.858
	PSNR	22.688	18.687	20.605	20.575	**24.224**
	FADE	0.938	0.810	1.096	**0.674**	0.700

The bold data is optimal data.

Table 6. Comparison table of unsupervised and semi-supervised wise defog algorithms.

Dataset	Metrics	USID [22]	DE [24]	SLA [25]	SDA-GAN [38]
I-haze (640 × 480)	Runtime	1.577	2.721	0.575	**0.240**
	SSIM	0.548	0.616	0.793	**0.860**
	PSNR	10.811	14.712	16.480	**18.039**
	FADE	**0.687**	1.546	1.277	0.739
O-haze (640 × 480)	Runtime	1.224	2.685	0.627	**0.221**
	SSIM	**0.800**	0.769	0.699	0.781
	PSNR	15.346	**17.454**	13.966	15.841
	FADE	**0.386**	0.981	0.658	0.387
SOTS-In	Runtime	1.139	2.380	0.351	**0.216**
	SSIM	0.808	0.833	**0.859**	0.848
	PSNR	17.666	19.207	**19.761**	18.768
	FADE	0.457	0.761	0.580	**0.447**
SOTS-Out	Runtime	1.010	2.100	**0.300**	0.361
	SSIM	0.730	0.857	**0.892**	0.841
	PSNR	14.424	19.166	**20.896**	18.332
	FADE	**0.460**	0.970	0.631	0.594

The bold data is optimal data.

5. Conclusions and Outlook

5.1. Conclusions

In this paper, we offer a comprehensive overview of recently developed defogging methodologies. We initially categorize defogging techniques into two broad groups: traditional approaches and deep learning-based methods. Traditional techniques encompass algorithmic methods that rely either on image enhancement techniques or on physical models of fog formation. On the other hand, deep learning-based defogging algorithms are further divided into network-based and method-based models. We then delve into each of these categories, highlighting well-known fog removal models, and outlining their core principles, strengths, limitations, and potential for future advancements.

5.2. Outlook

Despite the significant advancements in the field of image defogging, there remains considerable room for improvement. As a result, we point out several key issues that would inspire and catalyze future research. The issues are as follows:

1. **Computational efficiency and real-time processing**: Currently, there are a variety of applications that require real-time processing. Meanwhile, as network architectures grow much more complex, computational efficiency has become a critical concern. There is an urgent need for efficient defogging methods that reach a balance among the scale of parameters, inference time, as well as quantitative performance. Future efforts should be invested in the simplification of network architecture, possibly through model compression, quantization, and exploration of distributed computing, as well as hardware acceleration techniques.
2. **Evaluation metrics and benchmarks**: The majority of current defogging methods are evaluated using PSNR and SSIM, which necessitate paired images (both result and ground truth image) for assessment. However, for hazy images without corresponding clear images, developing robust evaluation methods is important. Thus, there is an urgent need to establish better non-reference evaluation metrics for image defogging.
3. **Adaptability across varying fog conditions**: Most defogging methods perform well under certain foggy conditions, but struggle with others. Future research should focus on developing adaptive methods that can dynamically adjust to different fog densities and atmospheric conditions without requiring retraining.
4. **Integration of physical models with deep learning**: Our paper discusses the advantages of both physical-model-based and deep learning-based defogging methods. Future work could benefit from hybrid models that integrate the physical understanding of light scattering in fog with the learning capabilities of neural networks. Such hybrid models may take advantage of both physical-model-based methods' efficiency and deep learning-based ones' accuracy.
5. **Unsupervised and self-supervised learning**: Given the scarcity of large, annotated datasets, unsupervised and self-supervised learning methods may be the best solution. Future research could explore the use of synthetic data generation, self-supervision, or other innovative methods to train robust defogging models without large-scale paired datasets.

Author Contributions: Software, T.L.; Investigation, Y.L.; Writing—original draft, M.S.; Writing—review & editing, J.Z.; Project administration, M.J. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

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

References

1. Muniyappan, S.; Allirani, A.; Saraswathi, S. A novel approach for image enhancement by using contrast limited adaptive histogram equalization method. In Proceedings of the 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, India, 4–6 July 2013; pp. 1–6. [CrossRef]
2. Shrivastava, A.; Jain, S. Single image dehazing based on one dimensional linear filtering and adoptive histogram equalization method. In Proceedings of the 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), Chennai, India, 3–5 March 2016; pp. 4074–4078. [CrossRef]
3. Thanh, L.T.; Thanh, D.N.H.; Hue, N.M.; Prasath, V.B.S. Single Image Dehazing Based on Adaptive Histogram Equalization and Linearization of Gamma Correction. In Proceedings of the 2019 25th Asia-Pacific Conference on Communications (APCC), Ho Chi Minh City, Vietnam, 6–8 November 2019; pp. 36–40. [CrossRef]
4. Weichao, H.; Zhi, Y.; Shangbin, J.; Ding, L. Research on Color Image Defogging Algorithm Based on MSR and CLAHE. In Proceedings of the 2020 Chinese Automation Congress (CAC), Shanghai, China, 6–8 November 2020; pp. 7301–7306. [CrossRef]

5. Liu, F.; Xue, Y.; Dou, X.; Li, Z. Low Illumination Image Enhancement Algorithm Combining Homomorphic Filtering and Retinex. In Proceedings of the 2021 International Conference on Wireless Communications and Smart Grid (ICWCSG), Hangzhou, China, 13–15 August 2021; pp. 241–245. [CrossRef]

6. Dai, W.; Ren, X. Defogging Algorithm for Road Environment Landscape Visual Image Based on Wavelet Transform. In Proceedings of the 2023 International Conference on Networking, Informatics and Computing (ICNETIC), Palermo, Italy, 29–31 May 2023; pp. 587–591. [CrossRef]

7. Zhang, R.; Zhao, F.; Dai, F.; Shi, Y. Algorithm for Fog-degraded Image Enhancement Based on Adaptive Fractional-order PDE. In Proceedings of the 2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP), Nanjing, China, 23–25 October 2020; pp. 129–135. [CrossRef]

8. Lei, L.; Wang, L.; Wu, J.; Bai, X.; Lv, P.; Wei, M. Research on Image Defogging Enhancement Technology Based on Retinex Algorithm. In Proceedings of the 2023 2nd International Conference on 3D Immersion, Interaction and Multi-Sensory Experiences (ICDIIME), Madrid, Spain, 27–29 June 2023; pp. 509–513. [CrossRef]

9. Liu, W.; Hou, X.; Duan, J.; Qiu, G. End-to-End Single Image Fog Removal Using Enhanced Cycle Consistent Adversarial Networks. *IEEE Trans. Image Process.* **2020**, *29*, 7819–7833. [CrossRef]

10. He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009. [CrossRef]

11. He, K.; Sun, J.; Tang, X. Guided Image Filtering. In *Computer Vision—ECCV 2010, Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 1–14. [CrossRef]

12. 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. [CrossRef] [PubMed]

13. Ju, M.; Ding, C.; Guo, Y.J.; Zhang, D. IDGCP: Image Dehazing Based on Gamma Correction Prior. *IEEE Trans. Image Process.* **2020**, *29*, 3104–3118. [CrossRef]

14. Schechner, Y.; Narasimhan, S.; Nayar, S. Instant dehazing of images using polarization. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), Kauai, HI, USA, 8–14 December 2001. [CrossRef]

15. Tan, R.T. Visibility in bad weather from a single image. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008. [CrossRef]

16. Fattal, R. Single image dehazing. *ACM Trans. Graph.* **2008**, *27*, 1–9. [CrossRef]

17. Tarel, J.P.; Hautière, N. Fast visibility restoration from a single color or gray level image. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 29 September–2 October 2009; pp. 2201–2208. [CrossRef]

18. Ju, M.; Ding, C.; Guo, C.A.; Ren, W.; Tao, D. IDRLP: Image Dehazing Using Region Line Prior. *IEEE Trans. Image Process.* **2021**, *30*, 9043–9057. [CrossRef]

19. Berman, D.; Treibitz, T.; Avidan, S. Non-local Image Dehazing. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 1674–1682. [CrossRef]

20. Oakley, J.; Satherley, B. Improving image quality in poor visibility conditions using a physical model for contrast degradation. *IEEE Trans. Image Process.* **1998**, *7*, 167–179. [CrossRef]

21. Kopf, J.; Neubert, B.; Chen, B.; Cohen, M.; Cohen-Or, D.; Deussen, O.; Uyttendaele, M.; Lischinski, D. Deep photo. *ACM Trans. Graph.* **2008**, *27*, 1–10. [CrossRef]

22. Golts, A.; Freedman, D.; Elad, M. Unsupervised Single Image Dehazing Using Dark Channel Prior Loss. *IEEE Trans. Image Process.* **2020**, *29*, 2692–2701. [CrossRef]

23. Shyam, P.; Yoon, K.J.; Kim, K.S. Towards Domain Invariant Single Image Dehazing. In Proceedings of the 2021 AAAI Conference on Artificial Intelligence, Online, 2–9 February 2021.

24. Golts, A.; Freedman, D.; Elad, M. Deep Energy: Task Driven Training of Deep Neural Networks. *IEEE J. Sel. Top. Signal Process.* **2021**, *15*, 324–338. [CrossRef]

25. Liang, Y.; Wang, B.; Zuo, W.; Liu, J.; Ren, W. Self-Supervised Learning and Adaptation for Single Image Dehazing. In Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, Vienna, Austria, 23–29 July 2022. [CrossRef]

26. 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]

27. Li, B.; Gou, Y.; Liu, J.; Zhou, J.T.; Peng, X. Zero-Shot Image Dehazing. *IEEE Trans. Image Process.* **2020**, *29*, 8457–8466. [CrossRef] [PubMed]

28. Ren, W.; Zhou, L.; Chen, J. Unsupervised single image dehazing with generative adversarial network. *Multimed. Syst.* **2023**, *29*, 2923–2933. [CrossRef]

29. Wang, W.; Wang, A.; Ai, Q.; Liu, C.; Liu, J. AAGAN: Enhanced Single Image Dehazing With Attention-to-Attention Generative Adversarial Network. *IEEE Access* **2019**, *7*, 173485–173498. [CrossRef]

30. Cong, X.; Gui, J.; Miao, K.C.; Zhang, J.; Wang, B.; Chen, P. Discrete Haze Level Dehazing Network. In Proceedings of the 28th ACM International Conference on Multimedia, New York, NY, USA, 12–16 October 2020; MM '20; pp. 1828–1836. [CrossRef]

31. Zheng, Y.; Zhan, J.; He, S.; Dong, J.; Du, Y. Curricular Contrastive Regularization for Physics-Aware Single Image Dehazing. In Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 17–24 June 2023; pp. 5785–5794. [CrossRef]

32. 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; pp. 16180–16189. [CrossRef]

33. Ren, W.; Liu, S.; Zhang, H.; Pan, J.S.; Cao, X.; Yang, M.H. Single Image Dehazing via Multi-Scale Convolutional Neural Networks. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016.

34. 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]

35. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 10–17 October 2021; pp. 9992–10002. [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. Li, L.; Dong, Y.; Ren, W.; Pan, J.; Gao, C.; Sang, N.; Yang, M.H. Semi-Supervised Image Dehazing. *IEEE Trans. Image Process.* **2020**, *29*, 2766–2779. [CrossRef]

38. Dong, Y.; Li, Y.; Dong, Q.; Zhang, H.; Chen, S. Semi-Supervised Domain Alignment Learning for Single Image Dehazing. *IEEE Trans. Cybern.* **2023**, *53*, 7238–7250. [CrossRef]

39. Liu, Y.; You, S.; Brown, M.; Tan, R. Haze Visibility Enhancement: A Survey and Quantitative Benchmarking. *Comput. Vis. Image Underst.* **2017**, *165*, 1–16. [CrossRef]

40. Xu, Y.; Wen, J.; Fei, L.; Zhang, Z. Review of Video and Image Defogging Algorithms and Related Studies on Image Restoration and Enhancement. *IEEE Access* **2016**, *4*, 165–188. [CrossRef]

41. Gui, J.; Cong, X.; Cao, Y.; Ren, W.; Zhang, J.; Zhang, J.; Tao, D. A Comprehensive Survey on Image Dehazing Based on Deep Learning. In Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, Montreal, QC, Canada, 19–27 August 2021. [CrossRef]

42. Ancuti, C.; Ancuti, C.O.; De Vleeschouwer, C. D-HAZY: A dataset to evaluate quantitatively dehazing algorithms. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016. [CrossRef]

43. Ju, M.; Ding, C.; Ren, W.; Yang, Y.; Zhang, D.; Guo, Y.J. IDE: Image Dehazing and Exposure Using an Enhanced Atmospheric Scattering Model. *IEEE Trans. Image Process.* **2021**, *30*, 2180–2192. [CrossRef]

44. 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] [PubMed]

45. Ullah, H.; Muhammad, K.; Irfan, M.; Anwar, S.; Sajjad, M.; Imran, A.S.; de Albuquerque, V.H.C. Light-DehazeNet: A novel lightweight CNN architecture for single image dehazing. *IEEE Trans. Image Process.* **2021**, *30*, 8968–8982. [CrossRef]

46. Engin, D.; Genc, A.; Ekenel, H.K. Cycle-Dehaze: Enhanced CycleGAN for Single Image Dehazing. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA, 18–22 June 2018. [CrossRef]

47. Yan, B.; Yang, Z.; Sun, H.; Wang, C. Ade-cyclegan: A detail enhanced image dehazing cyclegan network. *Sensors* **2023**, *23*, 3294. [CrossRef] [PubMed]

48. Targ, S.; Almeida, D.; Lyman, K. Resnet in resnet: Generalizing residual architectures. *arXiv* **2016**, arXiv:1603.08029.

49. Li, B.; Peng, X.; Wang, Z.; Xu, J.; Feng, D. An All-in-One Network for Dehazing and beyond. *arXiv* **2017**, arXiv:1707.06543.

50. Liu, X.; Ma, Y.; Shi, Z.; Chen, J. GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019. [CrossRef]

51. Xu, Q.; Wu, Z.; Yang, Y.; Zhang, L. The difference learning of hidden layer between autoencoder and variational autoencoder. In Proceedings of the 2017 29th Chinese Control And Decision Conference (CCDC), Chongqing, China, 28–30 May 2017. [CrossRef]

52. Wu, R.Q.; Duan, Z.P.; Guo, C.L.; Chai, Z.; Li, C. Ridcp: Revitalizing real image dehazing via high-quality codebook priors. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 18–22 June 2023; pp. 22282–22291.

53. Fan, C.N.; Zhang, F.Y. Homomorphic filtering based illumination normalization method for face recognition. *Pattern Recognit. Lett.* **2011**, *32*, 1468–1479. [CrossRef]

54. Spiegel, M.R. *Laplace Transforms*; McGraw-Hill: New York, NY, USA, 1965.

55. Ancuti, C.; Timofte, R.; Vleeschouwer, C. I-HAZE: A dehazing benchmark with real hazy and haze-free indoor images. In Proceedings of the Advanced Concepts for Intelligent Vision Systems: 19th International Conference, ACIVS 2018, Poitiers, France, 24–27 September 2018.

56. Ancuti, C.O.; Ancuti, C.; Timofte, R.; De Vleeschouwer, C. O-HAZE: A Dehazing Benchmark with Real Hazy and Haze-Free Outdoor Images. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA, 18–22 June 2018. [CrossRef]

57. 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]

58. Ancuti, C.O.; Ancuti, C.; Sbert, M.; Timofte, R. Dense-Haze: A Benchmark for Image Dehazing with Dense-Haze and Haze-Free Images. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019. [CrossRef]

59. Huynh-Thu, Q.; Ghanbari, M. Scope of validity of PSNR in image/video quality assessment. *Electron. Lett.* **2008**, *44*, 800–801. :20080522 [CrossRef]

60. Wang, Z.; Bovik, A.; Sheikh, H.; Simoncelli, E. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef]

61. Qiu, Y.; Zhang, K.; Wang, C.; Luo, W.; Li, H.; Jin, Z. MB-TaylorFormer: Multi-Branch Efficient Transformer Expanded by Taylor Formula for Image Dehazing. In Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France, 1–6 October 2023; pp. 12756–12767. [CrossRef]