



Fast no-reference deep image dehazing

Hongyi Qin¹ · Alexander G. Belyaev²

Received: 28 March 2024 / Revised: 5 July 2024 / Accepted: 7 August 2024 / Published online: 29 August 2024
© The Author(s) 2024

Abstract

This paper presents a deep learning method for image dehazing and clarification. The main advantages of the method are high computational speed and using unpaired image data for training. The method adapts the Zero-DCE approach (Li et al. in IEEE Trans Pattern Anal Mach Intell 44(8):4225–4238, 2021) for the image dehazing problem and uses high-order curves to adjust the dynamic range of images and achieve dehazing. Training the proposed dehazing neural network does not require paired hazy and clear datasets but instead utilizes a set of loss functions, assessing the quality of dehazed images to drive the training process. Experiments on a large number of real-world hazy images demonstrate that our proposed network effectively removes haze while preserving details and enhancing brightness. Furthermore, on an affordable GPU-equipped laptop, the processing speed can reach 1000 FPS for images with 2K resolution, making it highly suitable for real-time dehazing applications.

Keywords Image dehazing and clarification · No-reference neural network training · Real-time dehazing · Zero-DCE

1 Introduction

Bad weather image restoration is important for a number of real-world applications including video-based car driver assistance systems, autonomous drone navigation, and video surveillance. Haze, a prevalent atmospheric phenomenon, can substantially impair the effectiveness of high-level vision tasks, thereby underscoring the practical significance of a robust, well-generalized dehazing algorithm. A popular simplified model of images degraded by haze [1–3] is described by the equation:

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

Here $\mathbf{I}(x)$ and $\mathbf{J}(x)$ are the hazy and clear images, respectively. The term \mathbf{A} is the global atmospheric light. The transmission map $t(x) = e^{-\beta d(x)}$ quantifies the portion of

the light that reaches the camera, where $d(x)$ is the scene depth and β is the atmospheric scattering coefficient.

Image dehazing is an ill-posed problem. Traditionally, most prior-based algorithms and early learning-based methods have focused on estimating the transmission map $t(x)$ and global atmospheric light \mathbf{A} to reconstruct a clear image, as described by Eq. 1. More recently, advanced learning-based approaches have emerged that either directly predict the latent haze-free image or the residuals between haze-free and hazy images, thereby enhancing performance.

This paper introduces a novel deep learning method to attack the image dehazing problem. The approach follows the learning strategy of the Zero-DCE low-light image enhancement [4]: it is based on designing a multi-term no-reference loss function that distinguishes haze-free images from hazy ones. It doesn't require paired hazy and haze-free images of the same scenes and employs a lightweight neural network. As a result, in addition to haze suppression and image clarification, achieves a remarkably high image processing speed. The source code is available at <https://github.com/Hongyi311/Fast-No-Reference-Deep-Dehaze>.

1.1 Related work

As mentioned earlier, most dehazing methods can be divided into two categories. One is based on light scattered physical modeling [5–7] and hand-crafted image priors. Most of

✉ Hongyi Qin
hongyi.qin23@imperial.ac.uk

Alexander G. Belyaev
a.belyaev@hw.ac.uk

¹ Department of Electrical and Electronic Engineering, Imperial College London, London, UK

² Institute of Signals, Sensors and Systems, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, UK

these methods leverage priors to estimate the transmission map $t(x)$ and atmospheric light \mathbf{A} , subsequently obtaining the restored dehazed image $\mathbf{J}(x)$. The dark channel prior (DCP) [1] is pioneering work in this field, from which some improved methods based on DCP also emerged [8–10]. Further advancements have been achieved through the introduction of the Color Attenuation Prior [11], Gradient Channel Prior [12], Region Line Prior [13], and more recent innovations such as the Saturation Line Prior [14], Rank-One Prior [3], and Region Gradient Constraint Prior [15], etc., have further enhanced the performance of prior-based methods. These priors are often estimated based on statistical characteristics of haze-free and hazy images, boasting commendable interpretability and generalizability. However, their effectiveness may be diminished across images in varied scenarios.

The other family of methods exploits deep learning, aiming at restoring hazy images through well-trained neural network models. Cai et al. [2] first proposed a CNN-based model named DehazeNet to estimate the transmission map. Then Li et al. [16] proposed AODNet, which modified (1) to obtain recovered images in an end-to-end manner. Afterward, more sophisticated models emerged, such as incorporating attention mechanisms [17–20] and transformer-based architecture [21, 22], etc. They provided significant improvements in dehazing performance. However, such methods cannot overcome the reliance on paired clear images and synthesized hazy images as the training data. Real-world hazy scenes can exhibit a wide range of variability in terms of atmospheric conditions, lighting, scene content, etc. This poses an ongoing challenge for the model's generalization on real-world images. Although generative adversarial network (GAN) based methods [23, 24] have been proposed to eliminate the need for paired images, they still require careful selection of training data and usually incur significant training costs. UCL-dehaze [25] and C2PNet [26] introduced the novel unsupervised contrast learning approaches to dehazing, the models are extremely large in size and computationally intensive.

Recent progress in using deep learning models for image dehazing includes combining together extended haze models and sophisticated neural network architectures [27] and employing novel learning strategies [28–30]. Finally, it is worth mentioning recent works [31, 32] where multi-scale edge-aware image filters are proposed for image dehazing purposes.¹

Since dehazing algorithms often serve as data pre-processing for high-level tasks, real-time performance is one of the key objectives. In recent prior-based algorithms, Rank-One Prior [3] introduced GPU acceleration, achiev-

ing a processing time of 0.04 s per image for small-sized images (less than 720p). Region Gradient Constraint Prior [15] employs a parallel sliding window strategy, achieving impressive processing times of 0.004 s for a 512×512 image. Deep learning-based algorithms primarily enhance speed by reducing model size and FLOPs. With GPU acceleration, TOENet [19] can process a 2K hazy image in only 0.006 s. TSDNet [20] separates the dehazing task into three simple stages, achieving a processing speed of over 30 FPS with better quality.² It is difficult to decide which of these methods is the fastest one, as different hardware components and programming languages are used in the above-mentioned methods. However, it is safe to assert that the proposed method, achieving a processing speed of 1K FPS for 2K resolution images, is at least four times faster than the previously fastest method.

1.2 Contribution

In this paper, a zero-reference deep learning dehazing network is proposed. Inspired from Zero-DCE [4], the network does not rely on paired training data. The model's training is driven by a set of specially designed loss functions to evaluate the quality of the dehazed images. Additionally, rather than recovering clear images from hazy ones using the physical model described in (1), the approach from Zero-DCE is continued, applying high-order curves for pixel-wise adjustments on the dynamic range of the hazy image. Such a design enhances the model's generalization on real-world hazy images while simultaneously improving the recovery of details and brightness across various scenarios. Overall, the *contributions* of this study are as follows:

- The first image dehazing network that does not require paired or unpaired training data was developed, thereby improving the model's generalization on real-world hazy images.
- High-order curves were employed for pixel-wise adjustments to dehaze, rather than relying on physical models to recover images. This demonstrates the feasibility of using high-order curves for dynamic range adjustments in image dehazing.
- A comparison was made between the proposed method and several recently proposed image dehazing and clarification techniques. While it may not currently be classified as state-of-the-art (SOTA) in terms of natural image restoration, it performs exceptionally well in terms of fine image detail restoration and brightness enhancement. More importantly, the proposed method significantly out-

¹ We are grateful to one of the reviewers of the paper for pointing these two papers.

² The aforementioned speed performances are all sourced from the original papers.

performs others in terms of processing speed, making it highly efficient for real-time processing.

2 Proposed no-ref deep image dehazing ANN

In this section, the details of the proposed Fast No-reference Deep Image Dehazing network (FaNDID) are introduced, including the curve adjustments strategy, structure of the network, and zero-reference loss functions.

2.1 Network structure

Following the definition of the initial quadratic curve for image enhancement in Zero-DCE [4], our initial curve takes the following form:

$$D(x) = I(x) + \alpha(x)I(x)(1 - I(x)) \tag{2}$$

where $D(x)$ is the dehazed result of the input hazy image $I(x)$, $\alpha(x) \in [0, 1]$ is the curve parameters map with the same size as the input image, which means each pixel has its own curve to adjust the intensity. Then, iteratively applying this formula gives the high-order curve:

$$D_n(x) = D_{n-1}(x) + \alpha_n(x)(1 - D_{n-1}(x)) \tag{3}$$

where n is the iteration times. In our work, n is set to 8. This high-order curve enhances the adjustability of the dynamic range with its simple yet differentiable form.

The zero-reference dehazing network takes hazy images as input and outputs parameter maps for curve iteration. The network architecture follows the design used in Zero-DCE [4], as depicted in Fig. 1, consists of seven plain convolutional layers. In the first six layers, each employs 32 kernels of size 3×3 with a stride 1 and uses the ReLU activation function. The intermediate feature maps are symmetrically connected with skip concatenation. For a regular RGB image, eight iterations are performed when applying the high-order curve in Eq. 3 for each pixel in all three channels. Therefore, the final layer comprises 24 kernels, with the tanh activation function applied to limit the output range. This layer generates eight curve parameter maps for each color channel for iterative processing. The network is lightweight, with only 79K trainable parameters.

2.2 No-reference loss function

The loss function plays a pivotal role in zero-reference learning. Since our network is inspired by Zero-DCE, which specializes in enhancing low-light images, a naive approach was to invert the hazy image $\mathbf{I}(x, y)$ to treat $\mathbf{1} - \mathbf{I}(x, y)$

as a low-light image, subsequently enhancing its brightness before inverting the result once more. As demonstrated in Fig. 2, it's clear that this method yields unsatisfactory outcomes. However, the loss function in Zero-DCE has been proven to play a crucial role in balancing image brightness and color. Therefore, in addition to the four loss functions originally proposed in Zero-DCE, four additional functions are incorporated into the network to assess the quality of enhanced images, thereby driving improvements in dehazing performance. To ensure efficient training, these loss functions are all differentiable and possess relatively low computational complexity.

Dark Channel Loss The dark channel prior is defined as the local minimum of all pixels in a patch [1]:

$$D(x) = \min_{y \in \Omega_r(x)} \left(\min_{c \in \{r, g, b\}} I^c(y) \right) \tag{4}$$

where $\Omega_r(x)$ denotes a local patch centered at x , with size $r \times r$. Although it fails in sky and regions with bright colors, its simplicity and effectiveness make it a viable choice as one of the loss functions for measuring the degree of dehazing in enhanced images. Calculating the dark channel involves a process similar to a sliding window, which can significantly reduce computational speed. An effective approach is to directly divide the image into blocks with the same size as patches, then calculate the minimum pixel value in each block and assume that all the pixels in each block have the same dark channel value. $r = 8$ is used in this paper. Meanwhile, we dropped out those blocks with value larger than 0.7, which is usually the dark channel of the sky region. The loss function can be denoted as:

$$L_{dc} = \frac{W_1}{M} \sum_{i=1}^M \min_{x \in \Omega_i} x \tag{5}$$

where M is the number of blocks after dropping out and W_1 is the weight. L_{dc} provides an efficient approximation of the image's dark channel.

Contrast Loss Image contrast is related to total variation of the image gradient

$$C(I) = \sum_x |\nabla I(x)| \tag{6}$$

where ∇ denotes the gradient operator. Through multiple experiments, it has been found that using both the first and second derivatives of an image can enhance its edges. However, the former may lead to local overexposure and darkness, while the latter can refine texture and make the image brighter. Therefore, both of them are incorporated into a contrast-enhancing loss function L_{ce} and assign weights to balance their effects:

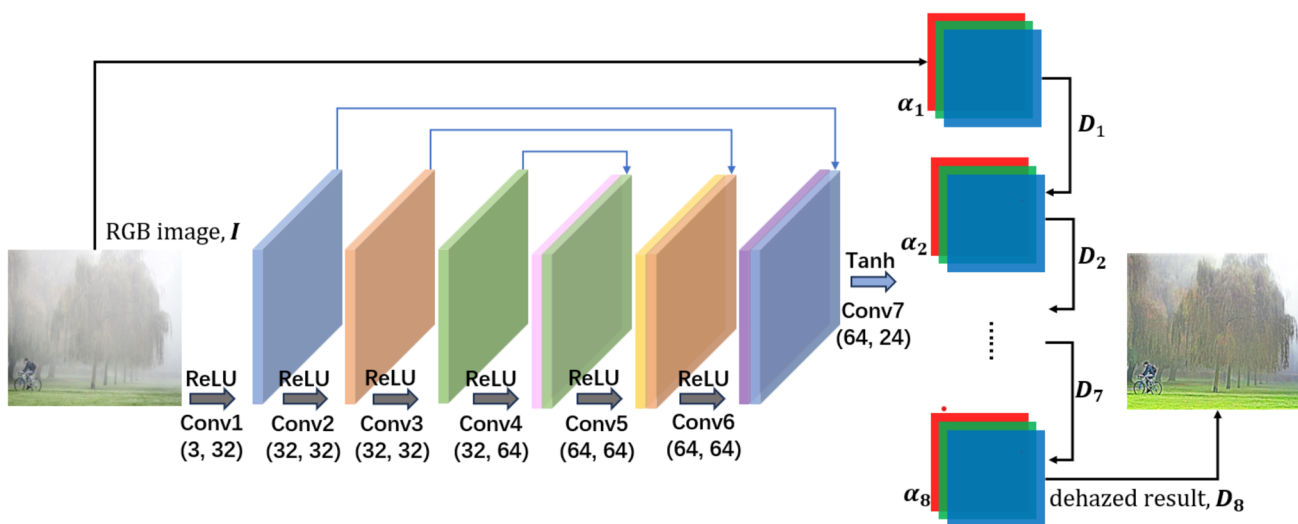


Fig. 1 Illustration of the structure representation and workflow of the network



Fig. 2 A naive approach to implement Zero-DCE on inverted hazy images. The right image in each pair is the final inverted low-light enhanced result

$$L_{ce} = -W_2 \ln \left[\frac{1}{N} \sum_x |\nabla I(x)| \right] - W_3 \ln \left[\frac{1}{N} \sum_x |\Delta I(x)| \right] \tag{7}$$

where N is the number of pixels of image \mathbf{I} , Δ stands for a 3×3 discrete Laplacian, and W_2 and W_3 are two weights to balance the influence between the ∇ and Δ components in the contrast loss function. Taking the negative natural logarithm allows it to enhance the contrast and texture details by reducing the loss.

Hue loss Hue disparity has been employed in [33] for detecting hazy areas, and it is defined as the difference in the hue channel between the original image I and the semi-inverse image I_{si} :

$$H(x) = \left| I^h(x) - I_{si}^h(x) \right| \tag{8}$$

where h means the hue channel, I_{si} is the maximum pixel-wise value between the original image and its inverse:

$$I_{si}(\mathbf{x}) = \max_{c \in (r, g, b)} [I^c(\mathbf{x}), 1 - I^c(\mathbf{x})] \tag{9}$$

Due to the high pixel intensities in regions heavily affected by haze, the semi-inverse image remains identical to the original one in these regions, resulting in low hue disparity. This concept can be used as a loss function. Theoretically, this

function can reduce the pixel values of hazy regions, but since it involves a comparison of hue channels, it also has an impact on the hue value of an image.

However, based on extensive validation with synthesized hazy and haze-free images, as well as the results obtained using state-of-the-art dehazing methods on real-world hazy images, it is observed that while the brightness and saturation of hazy images change significantly compared to the clear or dehazed images, the variation of hue is usually smaller, especially in areas with milder haze or in the presence of vividly colored objects within dense haze. Therefore, the research aims for an L_{hue} loss function component to predominantly impact the image brightness and saturation, while constraining its effect on hue: $L_{hue} =$

$$-W_4 \ln \left[\frac{1}{N} \sum_x \left| I^h(x) - I_{si}^h(x) \right| \right] + \frac{W_5}{N} \sum_x \left| I^h(x) - Y^h(x) \right| \tag{10}$$

where Y is the dehazed image, W_4 and W_5 are two positive weights.

Brightness loss To ensure that the bright regions do not become excessively dark, an additional function is introduced. This function calculates, separately for the original image and the enhanced image, the ratio of pixels with intensity greater than 0.7 to the total number of pixels and

compares these two ratios:

$$L_{bright} = W_6 \frac{B_{org} - B_{en}}{N} \quad (11)$$

Finally, the total loss is simply adding them together:

$$L_{total} = L_{dce} + L_{dc} + L_{ce} + L_{hue} + L_{bright} \quad (12)$$

where L_{dce} is the original loss function from Zero-DCE [4].

3 Experiments

3.1 Implementation details

Dataset. RESIDE- β [34] provides a large number of synthetic images with varying degrees of haze. From Part 1, 2000 images with different levels of haze were selected for training. Then, 500 outdoor images from Part 4 and 500 indoor images from SOTS (another subset) were used for testing. Additionally, to evaluate the dehazing performance on real-world hazy images, the testing set also included 1000 images from RTTS and 10 images from HSTS, both subsets of RESIDE- β [34], as well as 100 images from RUSH [3].

After numerous experiments, the following parameter settings were implemented: $W_1 = 0.2$, $W_2 = 0.3$, $W_3 = 0.5$, $W_4 = 0.1$, $W_5 = 5$, $W_6 = 5$, yielding the best visual results. The network optimization utilized the ADAM optimizer with default parameters and a fixed learning rate of $1e^{-4}$. The model was trained and tested using a laptop equipped with a 13th Gen Intel(R) Core(TM) i9-13900HX CPU @ 2.20 GHz with 32GB RAM and an Nvidia GeForce RTX 4060 GPU with 8GB VRAM.

3.2 Ablation study

3.2.1 Loss functions

Figure 3 illustrates the results of removing each loss function component. When L_{dc} is removed, some details within bright white backgrounds are lost, as L_{dc} originally serves to reduce the brightness in these areas, revealing details. Removing the first-order derivative component of L_{ce} leads to an overall brighter image but reduced contrast. On the other hand, $W_3 = 0$ makes the image much darker and loses numerous details. When the first part of L_{hue} is omitted, the colors in regions close to the camera become lighter, leading to an unclear appearance. When the second part restricting the hue is removed, the image undergoes severe color distortion. Finally, if L_{bright} is removed, the sky region becomes too dark. Therefore, each component of the loss function plays a crucial role in the visual quality of the image.

3.2.2 Number of layers in the network

Currently, the network utilizes 7 convolutional layers, resulting in rather high-order curves for adjusting the dynamic range of the image. Further testing was conducted with fewer layers, specifically 5 and 3 layers (since the network uses symmetric skip concatenation, it is preferable to reduce two layers at a time), and the dehazed results are shown in Fig. 4. Reducing the number of layers can further improve training speeds and alleviate the oversharpening issues. However, due to the reduction in the order of the adjustment curves, the dehazing capability of the network is weakened, performing poorly in those regions with heavy haze. In contrast, when the number of layers is 7, the details of distant buildings are relatively clear.

3.2.3 Dark channel patch size and sky threshold

When calculating the dark channel loss function, images are segmented using 8×8 patches. The loss function uses non-overlapping patches and averages the dark channel values of the patches after threshold filtering. Therefore, the patch size not only affects the dehazing details but also influences the overall color adjustment due to the varying number of patches. As shown in Fig. 5, if a 16×16 patch size is used, it means that more details might be obscured by the dark channel values. Additionally, fewer patches lead to a higher average loss value, making the dehazed image appear overly dim or color distorted, with white objects adjusted to the ambient color tone. On the other hand, reducing the patch size to 4×4 increases the number of patches, resulting in a lower initial loss value but poorer dehazing performance, with slightly sharpened images and deeper colors. The dark channels of images obtained with different patch sizes do not vary significantly in corresponding regions. Therefore, the number of patches significantly affects the initial loss value, impacting the dehazing results.

Additionally, the sky threshold is used in both the dark channel loss and brightness loss, serving to identify pixels or patches attributed to the sky region and thereby preserving its color integrity. Extensive testing established 0.7 as the optimal threshold. A lower threshold, such as 0.6, tends to obscure details in heavily hazed areas, while a higher threshold of 0.8 often results in images with darkened colors and reduced contrast, see Fig. 6. This effect is particularly pronounced in images where the sky appears grayish-white. In such cases, the pixel values in the sky region fall below the threshold, causing them to be treated as dense haze areas, which leads to excessive reduction in pixel values. Generally, a threshold of 0.7 is effective in preserving the natural color of the sky across various image characteristics.”

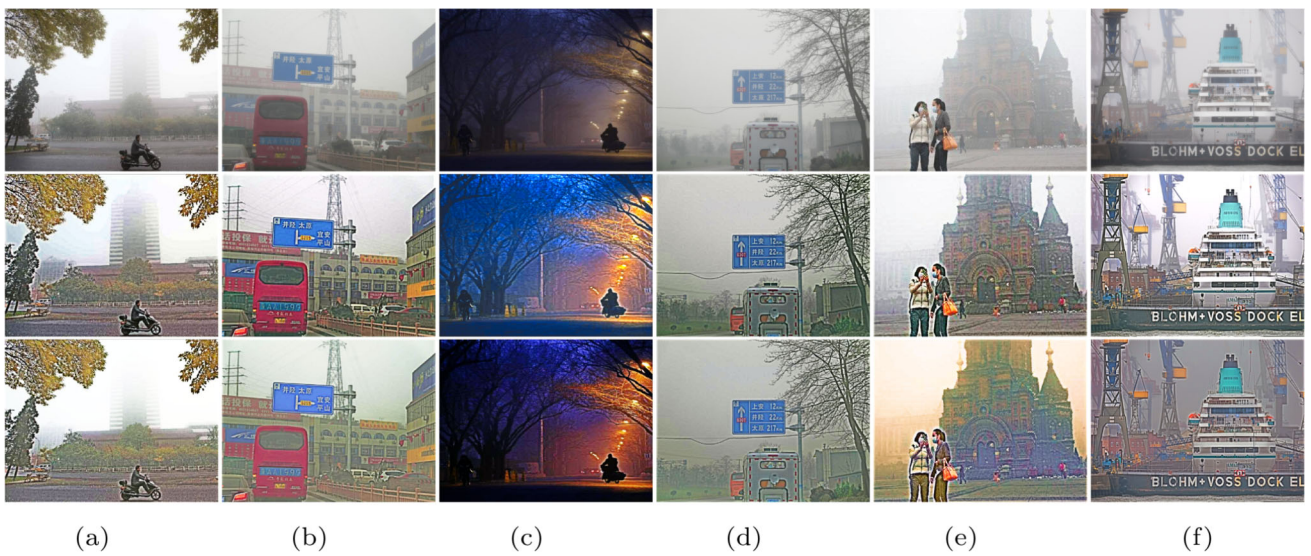


Fig. 3 Ablation study of the contribution of each component in the loss function. Top row: hazy images. Second row: de-hazed results when full weights applied. Bottom row: from left to right are results when **a** $W_1 = 0$, **b** $W_2 = 0$, **c** $W_3 = 0$, **d** $W_4 = 0$, **e** $W_5 = 0$, **f** $W_6 = 0$, respectively



Fig. 4 Ablation study on the number of network layers. **a** Hazy image. **b** 3 layers. **c** 5 layers. **d** 7 layers



Fig. 5 Ablation study on the patch size in dark channel loss. **a** is the hazy image. The patch size in **b** is 4×4 , in **c** is 8×8 , and in **d** is 16×16



Fig. 6 Ablation study on the value of sky threshold. The value in **a**, **d** is 0.6, in **b**, **e** is 0.7, and in **c**, **f** is 0.8. The corresponding hazy images can be found in Fig. 3. Zoom in for image details

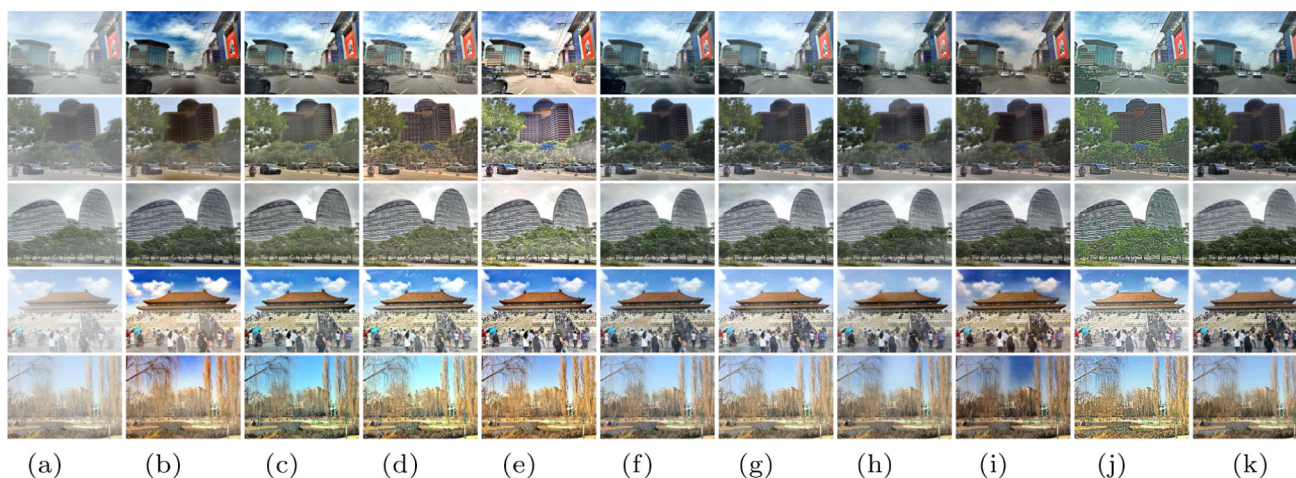


Fig. 7 Visual comparisons on outdoor synthetic hazy images from RESIDE- β . Images in **a** are hazy images, dehazed by **b** SLP, **c** ROP, **d** ROP+, **e** RGCP, **f** TOENet, **g** TSDNet, **h** DehazeFormer **i** C2PNet, and **j** the proposed FaNDID, respectively. **k** is GT. Zoom in for image details

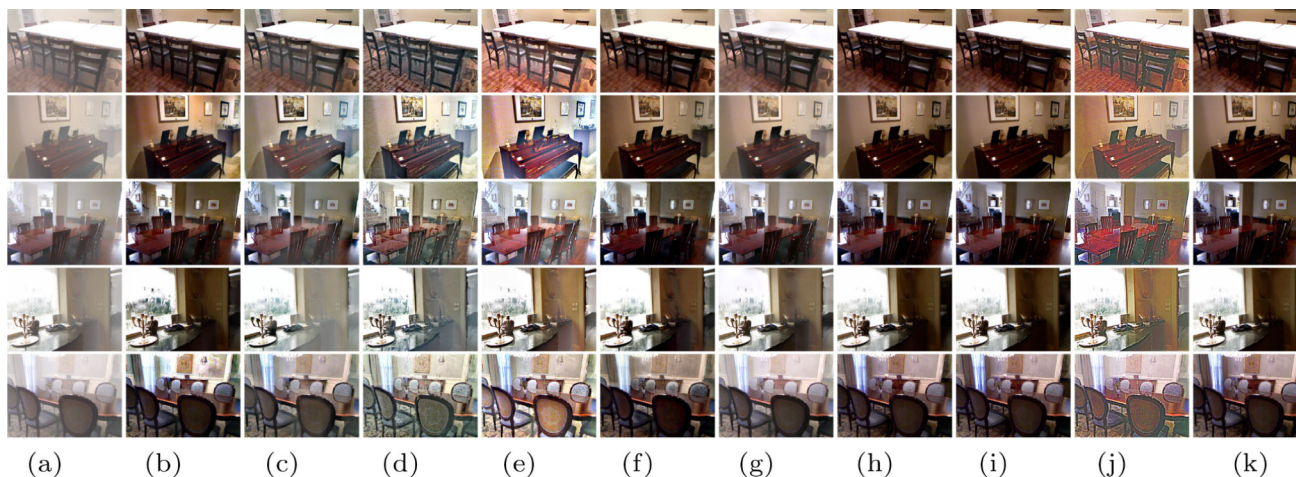


Fig. 8 Visual comparisons on indoor synthetic hazy images from RESIDE-SOTS. Images in **a** are hazy images, dehazed by **b** SLP, **c** ROP, **d** ROP+, **e** RGCP, **f** TOENet, **g** TSDNet, **h** DehazeFormer **i** C2PNet, and **j** the proposed FaNDID, respectively. **k** is GT. Zoom in for image details

3.3 Results and discussion

3.3.1 Qualitative comparison

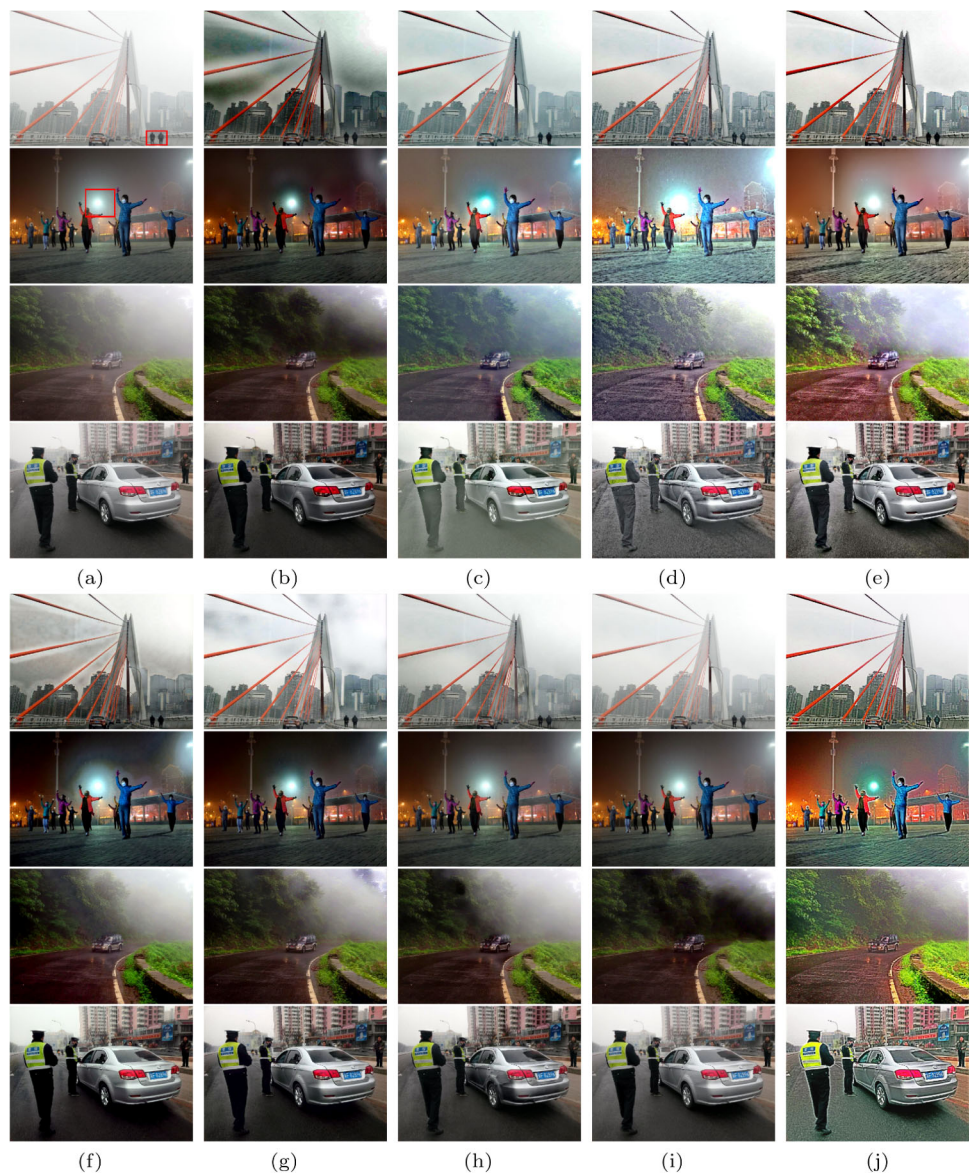
The dehazing performance of the proposed FaNDID on both synthetic and real-world hazy images was compared with those of state-of-the-art dehazing algorithms and models introduced in recent years. The prior-based methods include: Saturation Line Prior (SLP) [14], Rank One Prior (ROP and ROP+) [3], and Region Gradient Constrained Prior (RGCP) [15]. In terms of learning-based methods, comparisons were made with TOENet [19], TSDNet [20], DehazeFormer [22] and C2PNet [26].

Figures 7 and 8 illustrate the dehazing effects on outdoor synthetic hazy images from RESIDE- β [34] and indoor images from RESIDE-SOTS [34], respectively. It is evident that most learning-based methods, trained with paired

data to restore images to their original haze-free state, tend to achieve results that are closer to the ground truth compared to prior-based methods. Conversely, FaNDID employs an unpaired data training strategy, focusing on enhancing visual aspects such as brightness, contrast, and detail recovery. Despite some issues with oversaturation and sharpening, FaNDID's dehazing performance generally surpasses that of many prior-based methods.

However, synthetic hazy images usually differ significantly from real-world conditions due to the complex and varying nature of real-world haze, including factors like light scattering, particle size, and distribution. Therefore, the dehazing performance and generalization on real-world hazy images should receive more attention. FaNDID was tested on three real hazy image datasets: RESIDE-RTTS and RESIDE-HSTS (the realistic images subset), both from the work of Li

Fig. 9 Visual comparisons on real-world hazy images from RESIDE-RTTS. Images in **a** are hazy images, dehazed by **b** SLP, **c** ROP, **d** ROP+, **e** RGCP, **f** TOENet, **g** TSDNet, **h** DehazeFormer, **i** C2PNet and **j** the proposed FaNDID, respectively. Zoom in for image details



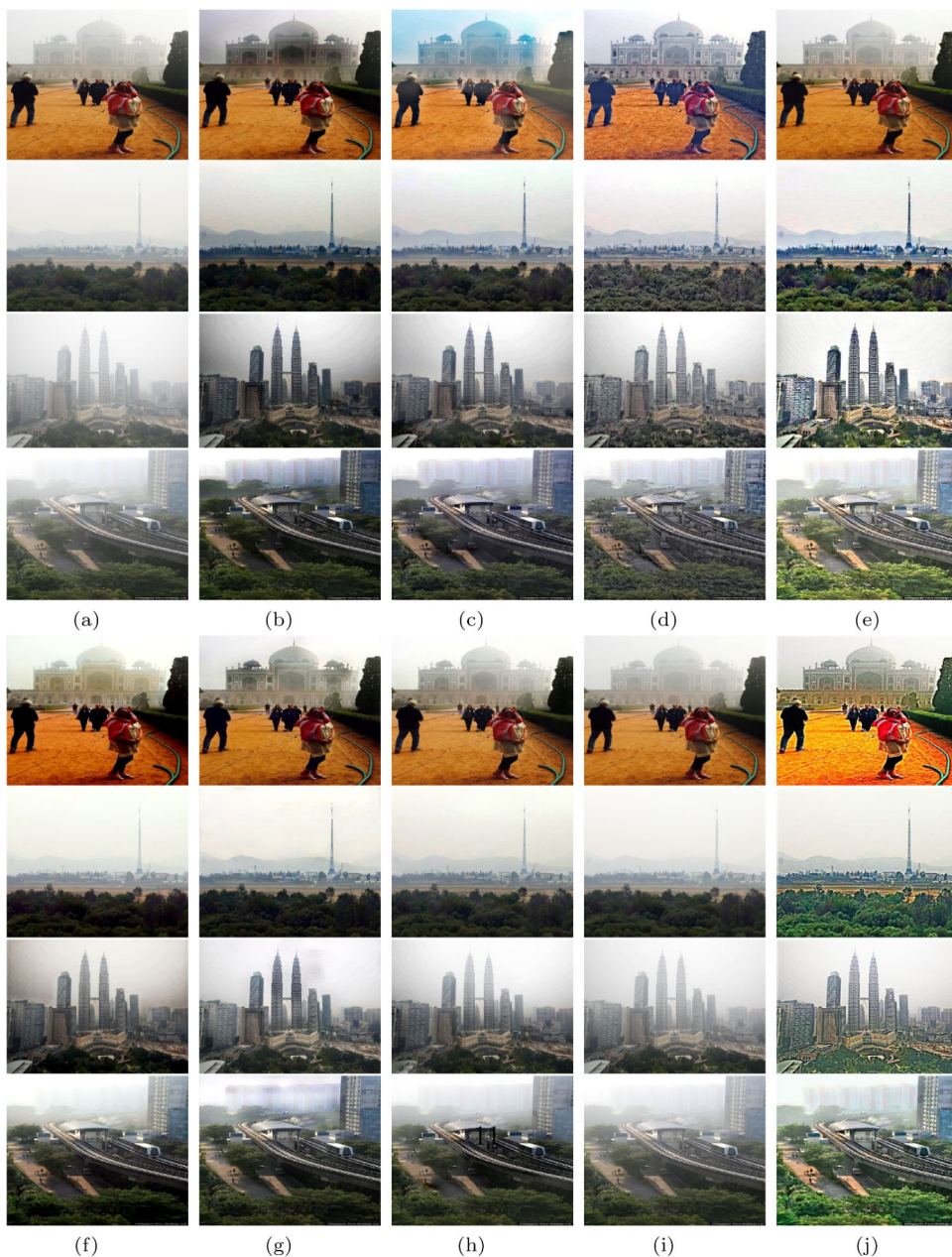
et al. [34], and RUSH [3]. The results are shown in Figs. 9, 10 and 11, respectively.

Observations indicate that among the evaluated dehazing methods, SLP [14] and all the other learning-based methods [19, 20, 22, 26] generally yield more natural-looking images. However, SLP and TOENet reduce image brightness significantly and introduce black shadows. Although TSDNet, DehazeFormer, and C2PNet perform better in dehazing without introducing black shadows, these methods do not always achieve good dehazing effects in all images, especially in low-light or heavy fog scenarios. Figure 10 is an example of this. On the other hand, ROP+ [35] and RGCP [15] excel in both dehazing efficiency and in enhancing image brightness. However, they also occasionally struggle with achieving complete haze removal and may destroy some details in the image. ROP [35], being the precursor to ROP+,

demonstrates this limitation more distinctly. Notably, in the images of the second row of Fig. 9, RGCP compromises the vibrancy of the colors in pedestrians' attire on the bridge. Meanwhile, in the third row, ROP+ tends to overemphasize the light sources, thereby masking the details of the surrounding environment. Similarly in Fig. 11, these three methods compromise the facial details of the motorcyclist in the first image, and due to excessive brightness enhancement, the background details of the forest in the third image are obscured.

Therefore, despite the proposed FaNDID introducing a mild degree of oversaturation and halo artifacts, it effectively maintains a harmonious equilibrium between image luminosity and detail preservation, offering a commendable solution in the context of dehazing challenges.

Fig. 10 Visual comparisons on real-world hazy images from RESIDE-HSTS. Images in **a** are hazy images, dehazed by **b** SLP, **c** ROP, **d** ROP+, **e** RGCP, **f** TOENet, **g** TSDNet, **h** DehazeFormer, **i** C2PNet and **j** the proposed FaNDID, respectively. Zoom in for image details



3.3.2 Quantitative comparison

Table 1 presents the referenced metrics, i.e. PSNR, SSIM, and multiscale SSIM, for dehazed results across 500 outdoor synthetic hazy images from the RESIDE- β dataset. As the qualitative analysis indicates, learning-based algorithms, which are trained with paired data, generally excel in restoring images closer to their original state, thereby achieving higher metrics scores. Conversely, prior-based methods, which rely on analyzing the statistical characteristics of images for dehazing, often do not perform as well on synthetic datasets due to the discrepancy between synthetic and real hazy images. FaNDID manages to bridge the gap

between learning-based and prior-based methods. It effectively reduces haze and achieves PSNR and SSIM scores that are higher than those of most prior-based methods, and the Multi-scale SSIM score is also competitive. This indicates the good performance of our method in preserving details.

Tables 2 and 3 apply no-reference image quality metrics NIQE [36], PIQE [37], and BRISQUE [38] to evaluate the quality of dehazed real-world images. Here, FaNDID outperforms most algorithms in terms of PIQE, while its performance in NIQE and BRISQUE is comparatively average. This is mainly due to the current results sometimes exhibiting issues of over-sharpening and over-saturation. However, these metrics should be considered as indicative rather than

Fig. 11 Visual comparisons on real-world hazy images from RUSH. Images in **a** are hazy images, dehazed by **b** SLP, **c** ROP, **d** ROP+, **e** RGCP, **f** TOENet, **g** TSDNet, **h** DehazeFormer, **i** C2PNet and **j** the proposed FaNDID, respectively. Zoom in for image details

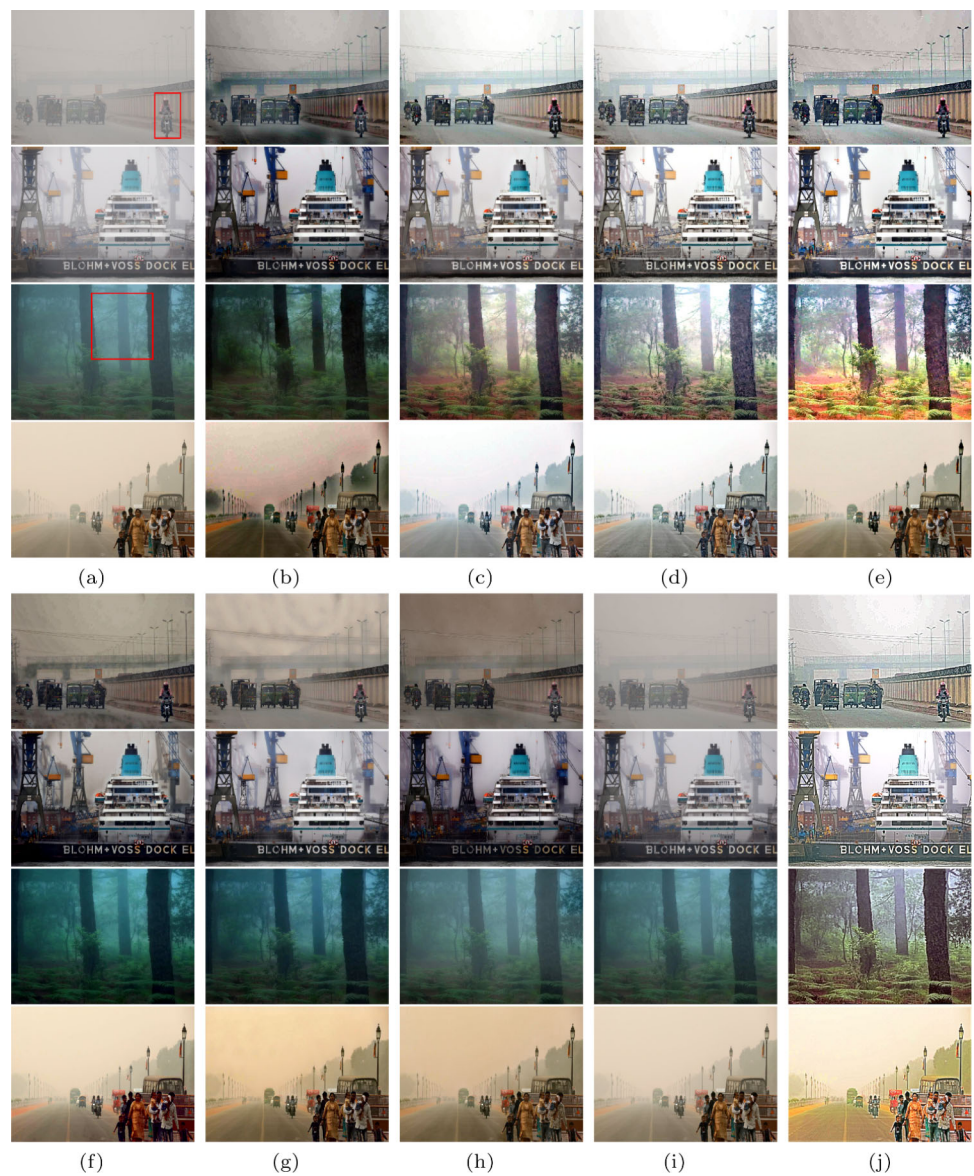


Table 1 The dehazed image quality assessment of 500 outdoor synthetic hazy images in RESIDE- β

Method	SLP	ROP	ROP+	RGCP	TOENet	TSDNet	Dehazeformer	C2PNet	FaNDID
PSNR \uparrow	17.08	15.13	16.02	14.69	22.01	18.94	20.58	21.52	16.42
SSIM \uparrow	0.76	0.69	0.71	0.65	0.89	0.88	0.88	0.89	0.77
Multi-SSIM \uparrow	0.88	0.82	0.86	0.83	0.95	0.96	0.95	0.96	0.83

Table 2 The dehazed image quality assessment of 1000 hazy images in RESIDE-RTTS

Method	SLP	ROP	ROP+	RGCP	TOENet	TSDNet	Dehazeformer	C2PNet	FaNDID
NIQE \downarrow	3.344	3.575	3.702	4.102	3.249	2.793	3.119	3.354	3.552
PIQE \downarrow	22.79	24.01	17.64	20.32	22.71	19.42	20.71	18.71	19.87
BRISQUE \downarrow	31.26	30.09	32.10	30.31	31.18	27.24	30.06	25.91	32.07

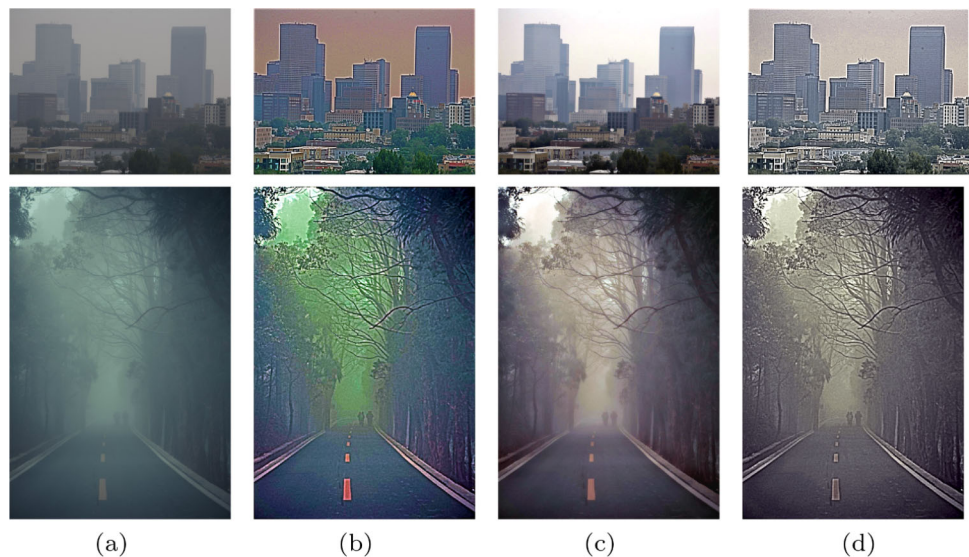
Table 3 The dehazed image quality assessment of 100 hazy images in RUSH

Method	SLP	ROP	ROP+	RGCP	TOENet	TSDNet	Dehazeformer	C2PNet	FaNDID
NIQE ↓	2.921	2.925	3.092	3.218	2.792	2.793	2.821	3.06	3.532
PIQE ↓	19.56	18.36	12.61	13.25	19.53	19.42	21.00	23.57	13.46
BRISQUE ↓	30.12	28.71	32.82	29.40	29.15	26.55	28.55	32.98	30.29

Table 4 Run-time (s) performance

Method	720p time/FLOPs	1080p time/FLOPs	2K time/FLOPs	Parameters	Platform
SLP	1.079	2.238	9.725	–	MATLAB
ROP	0.089	0.281	0.534	–	MATLAB
ROP+	0.182	0.363	1.611	–	MATLAB (GPU)
RGCP	0.047	0.136	0.269	–	MATLAB
TOENet	0.043/27.26G	0.048/61.33G	0.045/145.40G	273.6K	Pytorch (GPU)
TSDNet	0.006/338.49G	0.007/761.64G	0.007/1805.3G	1.22M	Pytorch (GPU)
Dehazeformer	0.012/172.77G	0.012/388.58G	0.013/692.05G	1.28M	Pytorch (GPU)
C2PNet	1.817/3251.4G	4.162/6537.2G	6.223/8322.7G	7.16M	Pytorch (GPU)
FaNDID	0.001/75.53G	0.0011/165.4G	0.0009/392.2G	79.2K	Pytorch (GPU)

Fig. 12 Sometimes our method may require a color amendment. **a** Hazy images. **b** Our method is applied. **c** ROP [35] nicely restores image colors but preserves a significant amount of haze. **d** Inter-image color transfer [39] is applied with the ROP images used as color donors



definitive assessments of dehazing quality and generalizability.

On the other hand, it is worth noting that the proposed method is exceptionally fast when utilizing GPU for acceleration, and the image size has almost no impact on processing speed. As listed in Table 4, it can achieve processing speeds of around 1000 FPS, making it highly suitable for real-time video dehazing. Compared to other standard deep learning methods that are trained on paired data, the model in this work is particularly lightweight, with only 79K trainable parameters, and its FLOPs are also highly competitive.

4 Limitations and future work

The proposed method is not free of drawbacks, as it may develop oversharpening effects including halo artifacts. This is also why it has not achieved the state-of-the-art (SOTA) status across all quality measures. The model aims to both dehaze and enhance the image, revealing many details in low-light conditions. However, this also results in higher overall image brightness, making the image appear less natural. This issue necessitates further optimization of the loss function components in future work. In addition, as illustrated in Fig. 12, these images have backgrounds with noticeable colors causing the network to accentuate these colors further.

This can be slightly alleviated by incorporating the ROP method [35] through an inter-image color transfer scheme [39]. Figure 12 shows the results using color transformation with the results of ROP as the source.

Currently, the loss function used in this model is too complicated but the authors failed to find a simpler loss function that delivers similar or better restoration quality. This will be a primary focus for future efforts to enhance this work. Another possible direction for future work consists of learning the loss function from examples.

The lightweight network with relatively shallow depth and simple architecture is used in this study, as seen in Fig. 1. Possibly employing a lightweight network with a more sophisticated architecture would lead to a better image restoration performance without sacrificing computational speed.

5 Conclusion

This paper introduces a zero-reference deep dehazing network that doesn't rely on paired images as a training dataset. It leverages several designed loss functions to evaluate the quality of dehazed images, driving the training process. Although it still exhibits some limitations, it outperforms other state-of-the-art methods in brightness enhancement and detail preservation. Furthermore, the network is lightweight and exceptionally fast, making it highly suitable for video dehazing applications. Promising future work includes video dehazing which requires adding temporal coherence terms to the loss function.

Acknowledgements The authors would like to thank the anonymous reviewers of this paper for their valuable and constructive comments.

Declarations

Conflict of interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(12), 2341–2353 (2010)
2. 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.* **25**(11), 5187–5198 (2016). <https://doi.org/10.1109/TIP.2016.2598681>
3. Liu, J., Liu, R.W., Sun, J., Zeng, T.: Rank-one prior: real-time scene recovery. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**(7), 8845–8860 (2023)
4. Li, C., Guo, C., Loy, C.C.: Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**(8), 4225–4238 (2021)
5. Narasimhan, S.G., Nayar, S.K.: Chromatic framework for vision in bad weather. In: *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662)*, vol. 1, pp. 598–6051 (2000). <https://doi.org/10.1109/CVPR.2000.855874>
6. Narasimhan, S.G., Nayar, S.K.: Vision and the atmosphere. *Int. J. Comput. Vis.* **48**, 233–254 (2002)
7. Fattal, R.: Single image dehazing. *ACM Trans. Gr.* **27**(3), 1–9 (2008)
8. Xu, H., Guo, J., Liu, Q., Ye, L.: Fast image dehazing using improved dark channel prior. In: *2012 IEEE International Conference on Information Science and Technology*, pp. 663–667. IEEE (2012)
9. Pan, J., Sun, D., Pfister, H., Yang, M.-H.: Blind image deblurring using dark channel prior. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1628–1636 (2016)
10. Peng, Y.-T., Cao, K., Cosman, P.C.: Generalization of the dark channel prior for single image restoration. *IEEE Trans. Image Process.* **27**(6), 2856–2868 (2018)
11. Zhu, Q., Mai, J., Shao, L.: A fast single image haze removal algorithm using color attenuation prior. *IEEE Trans. Image Process.* **24**(11), 3522–3533 (2015)
12. Singh, D., Kumar, V., Kaur, M.: Single image dehazing using gradient channel prior. *Appl. Intell.* **49**(12), 4276–4293 (2019)
13. Ju, M., Ding, C., Guo, C.A., Ren, W., Tao, D.: IDRLP: image dehazing using region line prior. *IEEE Trans. Image Process.* **30**, 9043–9057 (2021)
14. Ling, P., Chen, H., Tan, X., Jin, Y., Chen, E.: Single image dehazing using saturation line prior. *IEEE Trans. Image Process.* **32**, 3238 (2023)
15. Guo, Q., Zhang, Z., Zhou, M., Yue, H., Pu, H., Luo, J.: Image defogging based on regional gradient constrained prior. *ACM Trans. Multimed. Comput. Commun. Appl.* **20**(3), 1–17 (2023)
16. Li, B., Peng, X., Wang, Z., Xu, J., Feng, D.: AOD-Net: all-in-one dehazing network. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (2017)
17. Qin, X., Wang, Z., Bai, Y., Xie, X., Jia, H.: FFA-Net: feature fusion attention network for single image dehazing. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, pp. 11908–11915 (2020)
18. Liu, X., Ma, Y., Shi, Z., Chen, J.: GridDehazeNet: attention-based multi-scale network for image dehazing. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 7314–7323 (2019)
19. Gao, Y., Xu, W., Lu, Y.: Let you see in haze and sandstorm: two-in-one low-visibility enhancement network. *IEEE Trans. Instrum. Meas.* (2023). <https://doi.org/10.1109/TIM.2023.3304668>
20. Liu, R.W., Guo, Y., Lu, Y., Chui, K.T., Gupta, B.B.: Deep network-enabled haze visibility enhancement for visual IoT-driven

- intelligent transportation systems. *IEEE Trans. Ind. Inf.* **19**(2), 1581–1591 (2023). <https://doi.org/10.1109/TII.2022.3170594>
21. Guo, C.-L., Yan, Q., Anwar, S., Cong, R., Ren, W., Li, C.: Image dehazing transformer with transmission-aware 3d position embedding. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5812–5820 (2022)
 22. Song, Y., He, Z., Qian, H., Du, X.: Vision transformers for single image dehazing. *IEEE Trans. Image Process.* **32**, 1927–1941 (2023)
 23. Qu, Y., Chen, Y., Huang, J., Xie, Y.: Enhanced pix2pix dehazing network. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8160–8168 (2019)
 24. Dong, Y., Liu, Y., Zhang, H., Chen, S., Qiao, Y.: FD-GAN: generative adversarial networks with fusion-discriminator for single image dehazing. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, pp. 10729–10736 (2020)
 25. Wang, Y., Yan, X., Wang, F.L., Xie, H., Yang, W., Zhang, X.-P., Qin, J., Wei, M.: UCL-Dehaze: toward real-world image dehazing via unsupervised contrastive learning. *IEEE Trans. Image Process.* **33**, 1361–1374 (2024). <https://doi.org/10.1109/TIP.2024.3362153>
 26. Zheng, Y., Zhan, J., He, S., Dong, J., Du, Y.: Curricular contrastive regularization for physics-aware single image dehazing. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5785–5794 (2023)
 27. Feng, Y., Su, Z., Ma, L., Li, X., Liu, R., Zhou, F.: Bridging the gap between haze scenarios: a unified image dehazing model. *IEEE Trans. Circuits Syst. Video Technol.* (2024). <https://doi.org/10.1109/TCSVT.2024.3414677>
 28. Chen, Z., He, Z., Lu, Z.-M.: DEA-Net: single image dehazing based on detail-enhanced convolution and content-guided attention. *IEEE Trans. Image Process.* **33**, 1002 (2024)
 29. Cheng, D., Li, Y., Zhang, D., Wang, N., Sun, J., Gao, X.: Progressive negative enhancing contrastive learning for image dehazing and beyond. *IEEE Trans. Multimed.* (2024). <https://doi.org/10.1109/TMM.2024.3382493>
 30. Park, E., Yoo, J., Sim, J.-Y.: Universal dehazing via haze style transfer. *IEEE Trans. Circuits Syst. Video Technol.* (2024). <https://doi.org/10.1109/TCSVT.2024.3386738>
 31. Yadav, S.K., Sarawadekar, K.: An effective scale-aware edge-smoothing weighting constraint-based weighted guided image filter for single image dehazing. *Circuits Syst. Signal Process.* **42**(10), 6136–6159 (2023)
 32. Yadav, S.K., Sarawadekar, K.: Robust multi-scale weighting-based edge-smoothing filter for single image dehazing. *Pattern Recognit.* **149**, 110137 (2024)
 33. Ancuti, C.O., Ancuti, C., Hermans, C., Bekaert, P.: A fast semi-inverse approach to detect and remove the haze from a single image. In: *Asian Conference on Computer Vision*, pp. 501–514. Springer (2010)
 34. Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W., Wang, Z.: Benchmarking single-image dehazing and beyond. *IEEE Trans. Image Process.* **28**(1), 492–505 (2018)
 35. 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*, pp. 14802–14810 (2021)
 36. Mittal, A., Soundararajan, R., Bovik, A.C.: Making a “completely blind” image quality analyzer. *IEEE Signal Process. Lett.* **20**(3), 209–212 (2013). <https://doi.org/10.1109/LSP.2012.2227726>
 37. Venkatanath, N., Praneeth, D., Bh, M.C., Channappayya, S.S., Medasani, S.S.: Blind image quality evaluation using perception based features. In: *2015 Twenty First National Conference on Communications (NCC)*, pp. 1–6. IEEE (2015)
 38. Mittal, A., Moorthy, A.K., Bovik, A.C.: No-reference image quality assessment in the spatial domain. *IEEE Trans. Image Process.* **21**(12), 4695–4708 (2012). <https://doi.org/10.1109/TIP.2012.2214050>

39. Reinhard, E., Adhikhmin, M., Gooch, B., Shirley, P.: Color transfer between images. *IEEE Comput. Gr. Appl.* **21**(5), 34–41 (2001)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Mr. Hongyi Qin is an MSc student at Imperial College London, majoring in Communication and Signal Processing. He obtained his Bachelor's in Telecommunication Engineering from Xidian University in China and Heriot-Watt University in Scotland. His research interests are computer vision and image processing.



Alexander Belyaev is an associate professor at the School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, UK. His research interests include digital geometry processing, mathematical image analysis, applied partial differential equations, and recreational mathematics.