Pattern Recognition 165 (2025) 111615

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Self-supervised polarization image dehazing method via frequency domain generative adversarial networks

Rui Sun [a,b,c,*], Long Chen [a,b], Tanbin Liao [a,b], Zhiguo Fan [a,b]

[a] School of Computer Science and Information Engineering, Hefei University of Technology, No 485 Danxia Road, Hefei 230009, China
[b] Key Laboratory of Industry Safety and Emergency Technology, Hefei University of Technology, Hefei 230009, China
[c] Key Laboratory of Knowledge Engineering with Big Data, Ministry of Education of the Peoples Republic of China, Hefei 230009, China

ARTICLE INFO

Keywords:
Image dehazing
Polarization image
Frequency domain
Self-supervised
Generative adversarial network

ABSTRACT

Haze significantly hinders the application of autonomous driving, traffic surveillance, and remote sensing. Image dehazing serves as a key technology to enhance the clarity of images captured in hazy conditions. However, the lack of paired annotated training data significantly limits the performance of deep learning-based dehazing methods in real-world scenarios. In this work, we propose a self-supervised polarization image dehazing framework based on frequency domain generative adversarial networks. By incorporating a polarization calculation module into the generator, the Stokes parameters of airlight are accurately estimated, which are used to reconstruct the synthesized hazy image by combining the dehazed image generated via a densely connected encoder-decoder. Furthermore, we optimize the discriminator with frequency domain features extracted by frequency decomposition module and introduce a pseudo airlight coefficient supervision loss to enhance the self-supervised training. By discriminating between synthetic hazy images and real hazy images, we achieve adversarial training without the need for paired data. Simultaneously, supervised by the atmospheric scattering model, our network can iteratively generate more realistic dehazed images. Extensive experiments conducted on the constructed multi-view polarization datasets demonstrate that our method achieves state-of-the-art performance without requiring real-world ground truth.

1. Introduction

In modern society, where pollution is becoming increasingly severe, extreme weather conditions are occurring more frequently. The presence of haze particles in the atmosphere degrades the sharpness and contrast of images. Undoubtedly, haze significantly impairs the ability of computer vision algorithms to perceive scene information, leading to severe degradation or even failure in tasks such as recognition, detection, and segmentation. Therefore, research on improving visibility in hazy conditions, known as image dehazing [1], holds significant practical importance.

Existing dehazing methods are primarily categorized into prior-based and deep learning-based approaches. The theoretical foundation of prior-based methods lies in the atmospheric scattering model, which is widely used in computer vision to explain the physical principles behind imaging systems. By leveraging statistical priors to estimate parameters of the atmospheric scattering model, such as transmission and atmospheric light intensity, the dehazed image can be recovered. However, these methods [2–4] suffer from cascaded error propagation due to the reliance on handcrafted priors, limiting their effectiveness in improving visibility.

Unlike the methods that rely on a single hazy image, polarization-based dehazing methods [17–20] possess inherent advantages due to their utilization of polarization properties. The intensity of airlight varies regularly as the angle of polarization. By leveraging this property, the airlight intensity can be calculated using the degree of polarization, enabling the recovery of the dehazed image. Although the existing polarization-based dehazing methods effectively solve the problem of insufficient input information, they require multiple iterations to estimate parameters, resulting in low efficiency. Besides, as optimization-based approaches, they fail to fully exploit semantic and contextual information in image features to handle the spatially-variant real-world scattering.

In recent years, researchers have employed deep learning methods to estimate parameters in the atmospheric scattering model [5–8], which reduces the cascaded error to a certain extent. Additionally, many deep

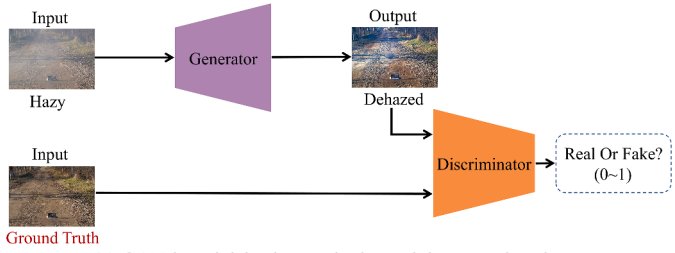
* Corresponding author.
E-mail address: sunrui@hfut.edu.cn (R. Sun).

https://doi.org/10.1016/j.patcog.2025.111615
Received 8 November 2024; Received in revised form 24 January 2025; Accepted 16 March 2025
Available online 17 March 2025

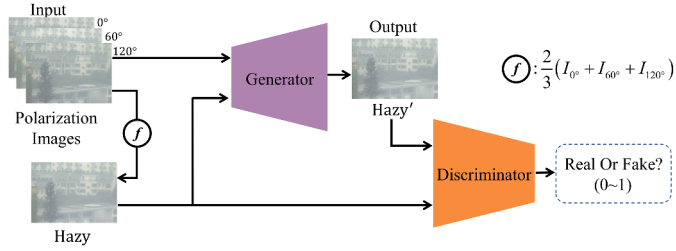
learning-based dehazing methods bypass the atmospheric scattering model [9–16,32–35], referred to as the end-to-end dehazing models. The features extracted by these models are more general, less time-consuming, more cost-effective, and exhibit the versatility of the dataset. As shown in Fig. 1(a), some GAN-based methods directly recover the dehazed image by learning the nonlinear mapping from hazy images to clear ones, but they are strictly limited by training datasets. Since it is impossible to simultaneously capture the real-world hazy image and its ground truth, synthetic datasets are predominantly used for training most networks. Consequently, their performances on real-world datasets is significantly inferior to that on synthetic datasets. Moreover, GAN-based methods are prone to generating artifacts, which adversely affects model training.

Leveraging the advantage of a large-scale polarization dataset [42] our team has previously collected, we aim to integrate polarization properties into deep learning-based dehazing methods to combine the strengths of both approaches and address their limitations. In this work, we propose a self-supervised polarization image dehazing method based on frequency domain generative adversarial networks (PGAN), hoping to enhance the performance and robustness of dehazing methods on real-world datasets.

First, polarization images are utilized to calculate the Stokes parameters of airlight (PGAN-A), while the dehazed image is generated using a densely connected encoder-decoder network (PGAN-J). The generated dehazed image and Stokes parameters are then employed to synthesize the hazy image through the physical model, serving as one of the input samples of the discriminator. By incorporating the polarization calculation module, our method requires only three hazy polarization images captured at different angles at a single time, eliminating the need for real-world ground truth. Second, we introduce a frequency decomposition module to optimize the discriminator based on the frequency domain distribution properties of hazy image. The input synthetic and original hazy images are separated in the frequency domain, and their respective high- and low-frequency sub-bands are combined as input samples for the discriminator, significantly enhancing its supervision capability.



(a) GAN-based dehazing methods need the ground truth.



(b) Our proposed self-supervised PGAN without need for the ground truth.

Fig. 1. Dehazing models use different architectures. (a) represents GAN-based dehazing methods, these methods rely on the ground truth of the hazy image while the ground truth of real scenes is usually difficult to define and obtain. And the synthetic data-driven methods have poor generalization in the real world. (b) represents our proposed self-supervised P-GAN, it does not require additional ground truth of the hazy image and effectively improves the authenticity of dehazed image by introducing the atmospheric scattering model. Our generated dehazed image is in the generator, the specific structure can be seen in Fig. 4.

Additionally, we design a pseudo airlight coefficient supervision loss to enhance the self-supervised training process. This loss function, grounded in the atmospheric scattering model, effectively mitigates the generation of excessive image noise and further improves the robustness and generalization capability of the network. By reformulating the problem as the discrimination between generated hazy image and real hazy image, our generator can produce realistic dehazed image that conforms to the atmospheric scattering model as much as possible. Our main contributions can be summarized as follows:

A novel self-supervised polarization image dehazing method is proposed, integrating polarization properties and frequency domain information within the generative adversarial learning framework.

A polarization-based generator is proposed, and the discriminator is optimized using the frequency distribution of hazy images. This approach addresses the limitations of paired training datasets, and significantly improves the dehazing performance of the network on real-world datasets.

A pseudo airlight coefficient supervision loss is designed for our self-supervised dehazing framework. Extensive qualitative and quantitative experiments demonstrate the superiority of our proposed method.

By introducing polarization supervision signals from the perspective of network architecture and loss function, generated images are constrained in the physical model, which not only improves the authenticity of generated images, but also ensures the timeliness of trained end-to-end dehazing model. The remainder of this paper is organized as follows. We first review the polarization image dehazing and the deep learning-based image dehazing in Section 2. Section 3 describes the proposed P-GAN model in detail, and Section 4 shows the experimental results. Finally, Section 5 concludes this paper.

2. Related works

2.1. Polarization-based image dehazing

As illustrated in Fig. 2, to advance the innovation and development of polarization dehazing algorithms, we constructed an atmospheric polarization information observation platform and UAV atmospheric polarization information observation platform, designed for horizontal and vertical observation perspectives, respectively. The fixed atmospheric polarization information observation platform integrates two intelligent polarization cameras (FGEA460 M and SZ-FGEB500C) from Hefei Shizhan Photoelectric Technology company, SONY Lucid polarization camera (IMX250MZR) and the related equipment of atmospheric environment monitoring (visibility meter, PM2.5 monitoring sensor, etc.). The polarization cameras adopted has the advantages of high extinction ratio and sensitivity, and has undergone strict polarization calibration [36]. Through this platform, we can better obtain

Fig. 2. Our fixed atmospheric polarization information observation platform, the UAV atmospheric polarization information observation platform can be seen in Section 4.1. We conducted extensive field observation experiments under foggy, hazy and other weather conditions, and obtained multi-target polarization data under various meteorological conditions.

polarization hazy data and explore the transmission characteristics of atmospheric polarized light in strong scattering media.

The atmospheric scattering model explains the contrast decay of hazy images from the perspective of the physical mechanism [1], as shown in Fig. 3. The mathematical representation of this model is as follows:

$$I(x) = D(x) + A(x) = J(x)t(x) + A_\infty(1 - t(x)) \tag{1}$$

where $I(x)$ is the captured hazy image, and x is the position of an image pixel. $I(x)$ consists the direct transmission $D(x)$ and airlight $A(x)$. J(x) is the object radiance, and it is the result to be recovered. A_∞ is the airlight radiance from the infinite distance.

Without considering the polarization properties of objects, early polarization-based dehazing methods controlled the received intensity of airlight by rotating a polaroid, utilizing the degree of polarization to describe the proportion of polarized light in natural light [17]. However, this method has a poor real-time performance. Currently, polarization cameras generally have focal focus planes that can acquire polarization images at multiple angles simultaneously. When polarization images at three or more angles (I0, I45, I90, I135) are available, the Stokes vector can be expressed as:

$$S = \begin{pmatrix} S_I \\ S_Q \\ S_U \\ S_V \end{pmatrix} = \begin{pmatrix} I_0 + I_{90} \\ I_0 - I_{90} \\ I_{45} - I_{135} \\ 0 \end{pmatrix} \tag{2}$$

where S_I denotes the total light intensity. S_Q and S_U denote the intensity of the linear polarization state. S_V denotes the intensity of circular polarization state, which is sufficiently small for it to be neglected ($S_V = 0$) in most natural light. Then the degree and angel of polarization noted as p and θ respectively can be calculated as follows:

$$p = \frac{\sqrt{S_Q^2 + S_U^2}}{S_I} \tag{3}$$

$$\theta = \frac{1}{2}\arctan\frac{S_U}{S_Q} \tag{4}$$

And the airlight radiance from the infinite distance can be estimated as follows:

$$A_\infty p_A = \frac{I_0 - A_\infty(1 - p_A)/2}{\cos^2\theta_A} \tag{5}$$

where p_A and θ_A denote the polarized degree and angel of airlight. Finally, the mathematical expression for obtaining the dehazed image J is given as follows:

$$J = \frac{S_I + S_Q/p}{1 + S_Q/(pA_\infty)} \tag{6}$$

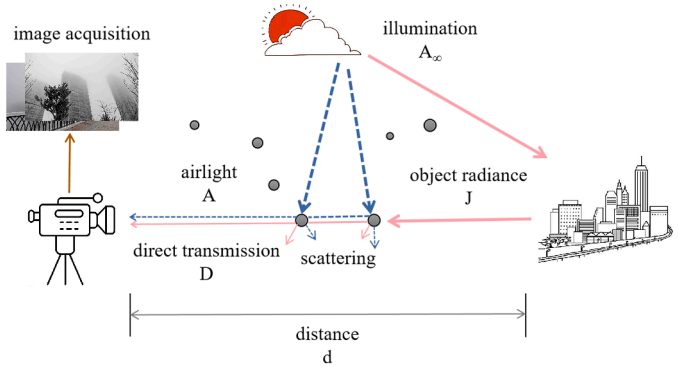


Fig. 3. Illustration of the atmospheric scattering model.

2.2. Deep learning-based image dehazing

Recently, researchers have utilized deep learning methods to perform image dehazing and achieved promising performance [35, 37–41]. Chen et al. [35] proposed detail-enhanced convolution and content-guided attention to boost the feature learning for improving the dehazing performance. Lyu et al. [38] introduced image priors in color spaces and proposed a multiple color space prior network to enhance the dehazing performance specifically for non-homogeneous hazy images. Unmanned aerial vehicle (UAV) dehazing has attracted much attention recently, Qiu et al. [40] present a novel UAV image dehazing framework to enhance perceptual tasks in foggy conditions. Sun et al. [41] proposed an unsupervised bidirectional contrastive reconstruction framework to enhance both the network's constraint and reconstruction capabilities.

In addition to the above methods, numerous GAN-based methods have been applied to image dehazing [9,21–24], which learn the mapping from hazy images to clear images without using the atmospheric scattering model. As shown in Fig. 1(a), the generator produces a dehazed image based on the input hazy image. Then, the discriminator evaluates the similarity between the dehazed image and ground truth. Typically, this similarity scores ranges from 0 to 1, with a higher score indicating greater resemblance to the ground truth. The generator improves the quality of dehazed images by utilizing this score as feedback. However, the training process requires a large number of paired datasets, which are extremely difficult to obtain in real-world scenarios. Moreover, the mapping from hazy images to clear ones lacks constraints imposed by the atmospheric scattering model. In contrast, our self-supervised PGAN eliminates the need for paired data and conforms to the constraints of the physical model, offering high real-world applicability and good interpretability, as shown in Fig. 1(b).

3. Proposed method

As illustrated in Fig. 4, the proposed PGAN mainly consists of a polarization-based generator (G) and a frequency-distribution-based discriminator (D). Unlike previous methods, our generator utilizes polarization images as input, and the output is a synthetic hazy image but not an intermediate generated dehazed image. Similarly, the input of our discriminator is modified to include both the synthetic hazy image and the original one. The original hazy image can be calculated from the polarization images. We optimize the loss function between the original hazy image and the synthesized hazy image to incorporate polarization supervision, thereby implementing a self-supervised training strategy.

3.1. The generator based on polarization dehazing

Our generator utilizes polarization images to calculate the Stokes vector of airlight and employs the atmospheric scattering model to re-synthesize the hazy image after generating the dehazed image, as shown in Fig. 4. PGN-J, constructed by a densely connected encoder-to-decoder, generates the dehazed image$J$from the original hazy image. Since each neuron in the densely connected layer receives input from all neurons in the previous layer, this architecture effectively facilitates feature extraction and reconstruction of scenes in low-level computer vision tasks [9], excelling at generating clearer dehazed images. The original hazy image I can be calculated from the polarization images using the Stokes vector as follows:

$$I = \frac{2}{3}(I_0 + I_{60} + I_{120}) \tag{7}$$

where I_0, I_{60}, I_{120} represent the polarization images at angle 0°, 60°, 120°, respectively.

PGAN-A is a polarization calculation module employed to calculate the Stokes vector of airlight from three polarization images. As illustrated in Fig. 5, this module utilizes the frequency distribution to

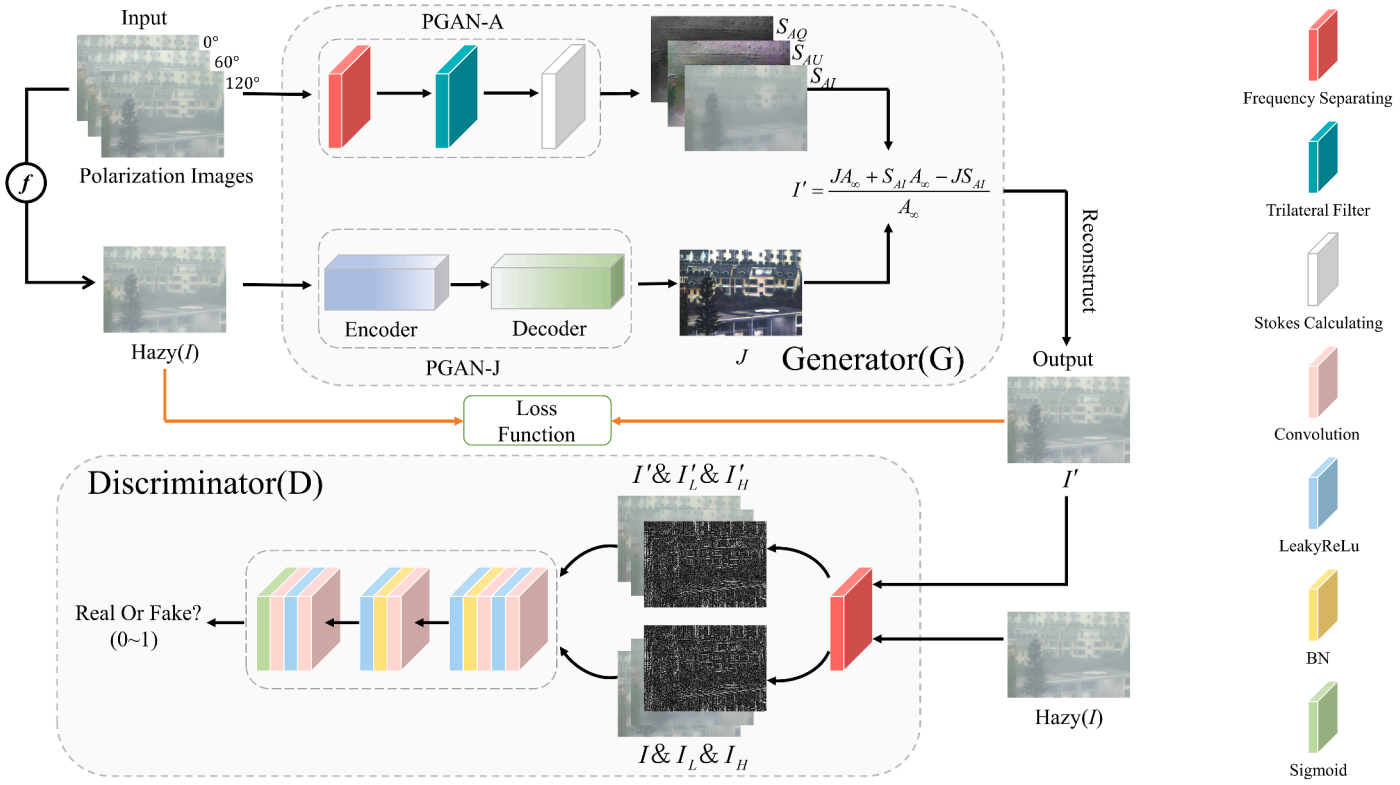


Fig. 4. The architecture of our proposed method PGAN. The generator input consists of three polarized images, and the original hazy image computed from f. f refer to (6). PGAN-J is a densely connected encoder-decoder used to generate the dehazed image, and PGAN-A is a polarization calculation module designed to estimate the atmospheric parameters in the haze image and provide physical constraints. The discriminator input consists of the original hazy image and the synthesized hazy image output by the generator. The discriminatory power of discriminators is enhanced by utilizing frequency distribution properties. For the trained model, we only use PGAN-J to generate dehazed images.

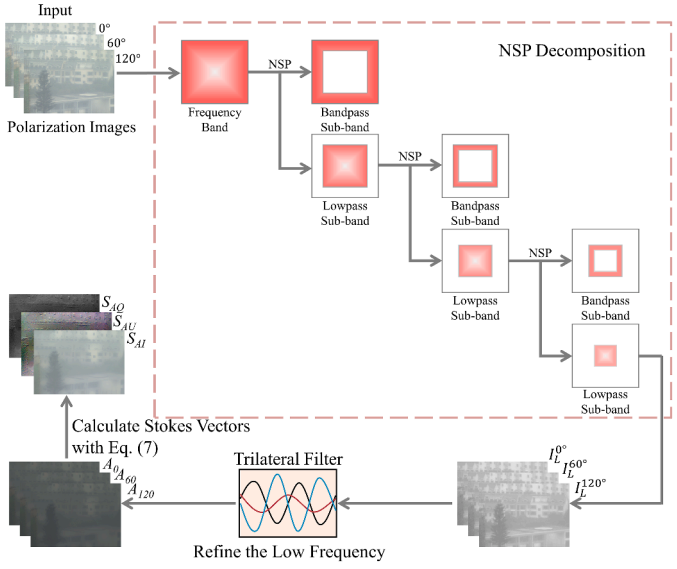


Fig. 5. The illustration of PGAN-A. PGAN-A is a polarization calculation module employed to calculate the Stokes vectors of airlight from three polarization images. We employ NSP decomposition to extract low frequency subbands from the polarization image, which are subsequently refined and used as airlight to calculate Stokes vectors.

separate airlight and avoid the halo effect, which does not need to discuss complex polarization properties of objects. Specifically, the polarization images are decomposed via the non-subsampled pyramid (NSP) [27].

$$H_n^{eq}(z) = \begin{cases} H_1\left(z^{2^{n-1}I}\right)\prod_{k=0}^{n-2} H_0\left(z^{2^k I}\right), 1 \le n \le k \\ \prod_{k=0}^{n-2} H_0\left(z^{2^k I}\right), n = k+1 \end{cases} \tag{8}$$

Where $H_n^{eq}(z)$ is the cascade filter at the n-th stage, $H_0(z)$ representing a low-pass decomposition filter and $H_1(z)$ is a high pass decomposition filter. It can be found that NSP decomposes a low-pass subband and a band-pass subband for the first time. As the number of decompositions increases, the low-pass subband will be decomposed many times, and finally k band-pass subbands and a low-pass subband will be generated in total. Follow our previous work [20], we set the number of decomposition levels to 4 to obtain the optimal decomposition image. The low-frequency subbands obtained by NSP decomposition are denoted as $I_L^{0°}$, $I_L^{60°}$, $I_L^{120°}$. Because of the airlight constraints, the decomposed low-pass sub-bands which are used as airlight in different polarization angles should be refined. We perform the trilateral filtering to refine the low-frequency subbands.

$$\begin{pmatrix} A_k = \max\left(\min\left(B(x), I^k(x)\right), 0\right) \\ B(x) = I_{L'}^{k^*}(x) - median_{sv}\left(\left|I^k(x) - I_{L'}^{k^*}(x)\right|\right) \\ I_{L'}^{k^*}(x) = median_{sv}\left(I^k(x)\right) \\ k = 0, 60, 120 \end{pmatrix} \tag{9}$$

Where $median_{sv}$ represents the median filter, A_0, A_{60}, A_{120} are the airlight of three polarization hazy images at angle 0°, 60°, 120°, respectively. With the airlight, the Stokes vector can be calculated as follows:

$$S_A = \begin{pmatrix} S_{AI} \\ S_{AQ} \\ S_{AU} \end{pmatrix} = \begin{pmatrix} \frac{2}{3}(A_0 + A_{60} + A_{120}) \\ \frac{2}{3}(2A_0 - A_{60} - A_{120}) \\ \frac{2}{\sqrt{3}}(A_0 - A_{120}) \end{pmatrix} \tag{10}$$

On this basis, we utilize Eq. (4) and (5) are to calculate A∞. Lastly, the synthetic hazy image I' can be calculated as follows:

$$I' = \frac{JA_\infty + S_{AI}A_\infty - JS_{AI}}{A_\infty} \tag{11}$$

Discussion. The polarization dehazing method has its own unique advantages compared to other dehazing methods, as it alleviates the fundamental flaw of dehazing tasks - insufficient acquisition of scene information. This advantage is specifically reflected in the ability to calculate the original fog image and polarization degree from the polarized image, and then estimate the physical model, which is very effective in real-world dehazing. At present, although deep learning methods can achieve good performance through style transfer [31] or layer disentanglement [32], they ignore the application of polarization characteristics. We introduce polarization characteristics on the basis of GAN-based dehazing method, improve the performance and robustness of the network on real data through the introduction of polarization as a physical mechanism, solve the limitations of training data, and enhance its interpretability, thus making efforts to promote the application of polarization characteristics in other deep learning methods.

3.2. The discriminator based on frequency distribution

As illustrated in Fig. 4, our discriminator separates the input original hazy image I and synthetic one I'in the frequency domain, extracting their high-frequency and low-frequency sub-bands, which combined with their corresponding hazy images as samples of the discriminator. To train the network, we utilize the dehazed image J as an intermediate result of the generator, aiming for the discriminator to guide the generator in producing a synthetic hazy image that closely resembles the original one. This process ultimately yields a clearer dehazed image under the physical constraints of the atmospheric scattering model. Furthermore, to enhance the discriminator's input diversity, we incorporate additional training constraints (frequency information). Direct transmission and airlight are distributed in the high-frequency IH and low-frequency IL sub-bands of the image frequency band [20], respectively. Therefore, the low-frequency components of hazy images are richer than clear ones due to the higher proportion of airlight. This characteristic is leveraged to optimize the design of our discriminator by integrating a frequency decomposition module.

To jointly optimize model's dehazing capability, it should be noted that the frequency decomposition method used in this module is consistent with PGAN-A as shown in Fig. 4. Since low-frequency sub-bands IL can be used as airlight after constraints, the supervised learning of them can actually be regarded as the supervised learning of airlight coefficient, which inspired us to design the supervised loss function of the pseudo-airlight coefficient. In previous work, researchers typically used generated dehazed images and ground truths as paired learning samples for the discriminator, with less exploration of the model's supervisory information. The discriminator we designed can more effectively distinguish between real and synthetic data, thereby generating more realistic and satisfactory dehazed images. We conducted a detailed experimental analysis on the effectiveness of the frequency domain decomposition module in Section 4.4.

3.3. The design of loss functions

This subsection focuses on the design of the loss function in our network. In addition to incorporating common losses such as pixel-level loss, SSIM loss, and adversarial loss into our framework, we also introduce a pseudo airlight coefficient supervision loss. This loss considers the low-frequency components of the discriminator input and the application of the physical model. Since low-frequency components represent the primary information of the image, designing a pseudo atmospheric scattering coefficient supervision loss based on these components enhances the constraints on network training, resulting in a more robust model.

3.3.1. Pixel-level loss

The pixel-level loss is utilized to measure the fidelity between the real and fake sample. In our method, given an original hazy image I, the synthetic hazy image output by the generator is I'. Then the pixel-level loss function LP in the form of L1 on N samples can be calculated as follows:

$$L_P = \sum_{i=1}^{N} \| I'_i - I_i \| \tag{12}$$

3.3.2. SSIM loss

The Structure Similarity Index Measure (SSIM) is an important reference image quality evaluation index [25], which comprehensively considers differences in brightness, contrast and structure. It accurately reflects the image quality of human perception so as widely used in the algorithm performance evaluation of computer vision tasks. In our method, SSIM can be calculated as follows:

$$SSIM(I', I) = \frac{(2\mu_{I'}\mu_I + C_1)(2\sigma_{I'I} + C_2)}{(\mu_{I'}^2 + \mu_I^2 + C_1)(\sigma_{I'}^2 + \sigma_I^2 + C_2)} \tag{13}$$

where$\mu$and $\sigma^2$are the average value and the variance, respectively. $\sigma$is the covariance. C_1,$C_2$are constants used to maintain stability. SSIM ranges from 0 to 1 and the SSIM loss is defined as follows:

$$L_{SSIM} = 1 - SSIM(I', I) \tag{14}$$

3.3.3. Adversarial loss

The adversarial loss encourages the generator to continuously improve the sample quality by minimizing the differences between generated and real samples, while maximizing the ability of the discriminator. We integrate the adversarial mechanism into the training process, enabling the model to learn more effective representations and features, thereby achieving better performance and robustness. In our method, the hazy images combined with high and low sub-bands samples are expressed as$I \cup I_H \cup I_L$, then the adversarial loss can be expressed as follows:

$$L_{Adv} = \log(1 - D(G(I \cup I_H \cup I_L))) \tag{15}$$

3.3.4. Pseudo airlight coefficient supervision loss

Airlight is closely related to the hazy image through the atmospheric scattering model, which directly affects the brightness and the distribution of fog. Therefore, the more similar the synthetic hazy image of generator is to the real hazy image, the more similar the dehazing image generated by PGAN-J is to the real clear image. Accurate airlight estimation can significantly improve the defogging effect, make the recovered image clearer and reduce artifacts. We expect to penalize differences of airlight between the original and synthetic hazy images in the network by designing the pseudo airlight coefficient supervision loss. Arlight can be estimated from the low-frequency sub-band after frequency decomposition in the discriminator. We regard the low-frequency subbands of the synthetic and original hazy images constrained by Eq. (8) as airlight, denoted as$A'$and A, respectively. Following the pixel-level loss, we force airlight to converge by MAE loss. Then the pseudo airlight coefficient supervision loss on the N samples is defined as

$$L_A = \sum_{i=1}^{N} \| A'_i - A_i \| \tag{16}$$

Finally, the loss function of the whole network consists of the pixel-level loss, SSIM loss, adversarial loss, and pseudo airlight coefficient supervision loss, and it can be expressed as

$$L = \lambda_P L_P + \lambda_{SSIM} L_{SSIM} + \lambda_{Adv} L_{Adv} + \lambda_A L_A \tag{17}$$

whereλrepresent the weight coefficient of the corresponding loss.

4. Experimental results

4.1. Datasets construction and parameters settings

Dataset. Most of the existing hazy image datasets are synthetic, which often fail to accurately reflect real-world haze distribution, particularly under strong wind conditions. Additionally, there are also many problems in a limited number of real-world datasets which creating artificial haze using a smoking generator. For a long time, the disclosure of datasets in the field of polarization dehazing has not attracted enough attention from the research community, which has severely hindered the widespread adoption and application of polarization methods.

In order to efficiently collect polarization data, we built different polarization data collection platforms with polarization cameras, and carried out long-term data collection work from November 2021. We have collected >30,000 hazy polarization images, and some of them have been open in the form of datasets, as shown in Fig. 6. We selected 600 polarization images from each of three different haze concentrations in the horizontal direction. The haze concentration of image was based on the value of no reference image quality index FADE [26] and visibly: thin haze (0 < FADE≤6.7, 1km< visibility ≤3 km), heavy haze (6.7 < FADE≤7.9, 200 m < visibility ≤1 km), and dense haze (7.9 < FADE≤9.0, visibility ≤200 m). In addition to the horizontal data, vertical data also holds significant experimental value, as it enables the evaluation of dehazing methods in aerial scenarios. Therefore, we also tilized 600 polarization images captured vertical direction using our unmanned aerial vehicle (UAV) platform, as illustrated in Fig. 7.

Our datasets do not explicitly use different scene names as classification criteria to reflect data diversity. Instead, different objects and backgrounds in the large view field of data acquisition equipment (masonry buildings, metal windows and tarmac pavement, etc.,) are used to reflect the diversity. We can evaluate the effects of different degree of scattering degradation on scenes with different polarization characteristics in the same view field, and the restoration effects of

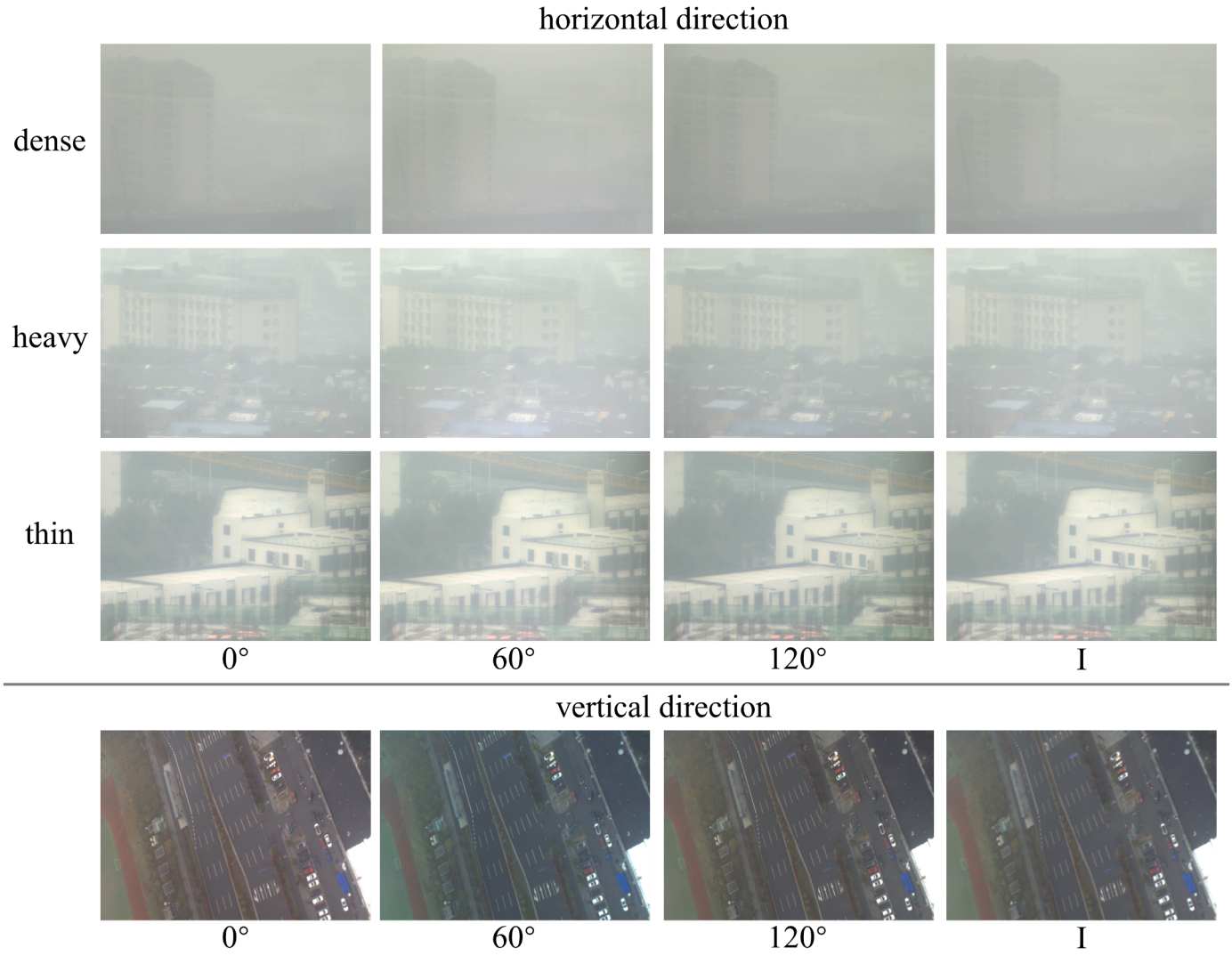


Fig. 6. Example images of our constructed dataset. We collect color images with polarization angles of 0°, 60°, and 120°, as well as color hazy image I. The image resolution is 2480×1860.The obtained hazy images are divided into three concentration levels: dense, heavy, and thin. In addition, the dataset covers both horizontal and vertical directions, effectively solving the problem of a single perspective in existing datasets.

Fig. 7. The UAV for vertical direction data acquisition.

related algorithms on various targets and backgrounds with different polarization characteristics.

Parameters. We set the four weight parameters λ_P, λ_{SSIM}, λ_{Adv}, λ_A to 1, 0.5, 0.1 and 1, respectively. The Nonsubsampled Pyramid (NSP) was used in all frequency domain decomposition modules, and the layer of decomposition was set to 4. In terms of the size of the input image, the original size captured by the polarization camera is 2480×1860×3. The polarization image and the original fog image have been downsampled, and the downsampled image size is 320×240×3. The training of the entire network is conducted on one Nvidia RTX A6000 GPU.

4.2. Qualitative experimental results in the horizontal and vertical directions

We chose dehazing algorithms based on priors and polarization including DCP [2] and PBD [17]. We also chose two dehazing algorithms based on deep learning including YOLY [32], D4 [24] and Dehamer [11], DehazeFormer [33] and DEA-Net [35]. Our method is defined as

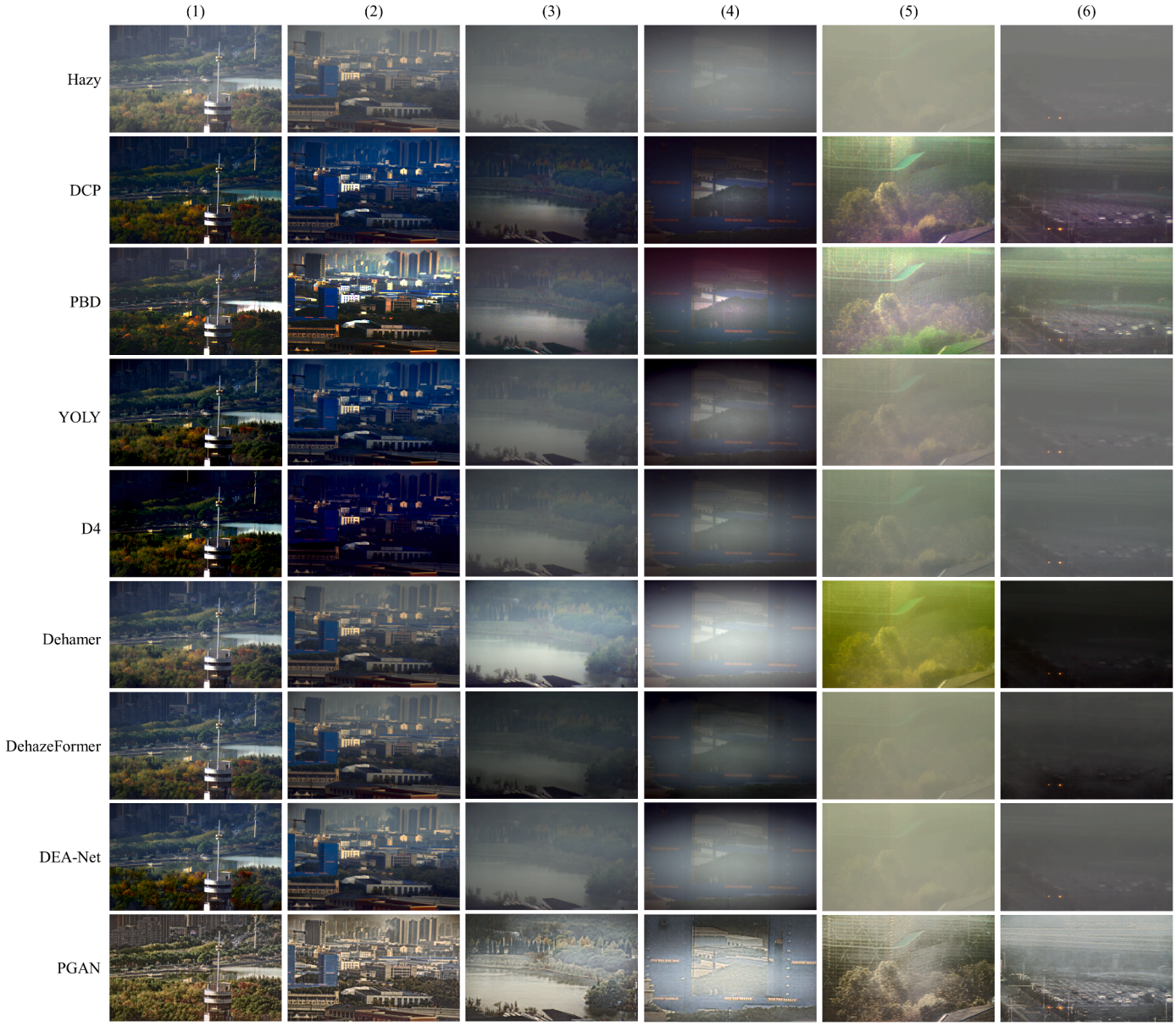


Fig. 8. Qualitative experimental results of the data in horizontal direction. It can be observed that our proposed PGAN demonstrates a better dehazing effect than other polarization-based and deep-learning-based state-of-the-art methods. Especially in the dense haze scenes, PGAN presents a strong dehazing ability.

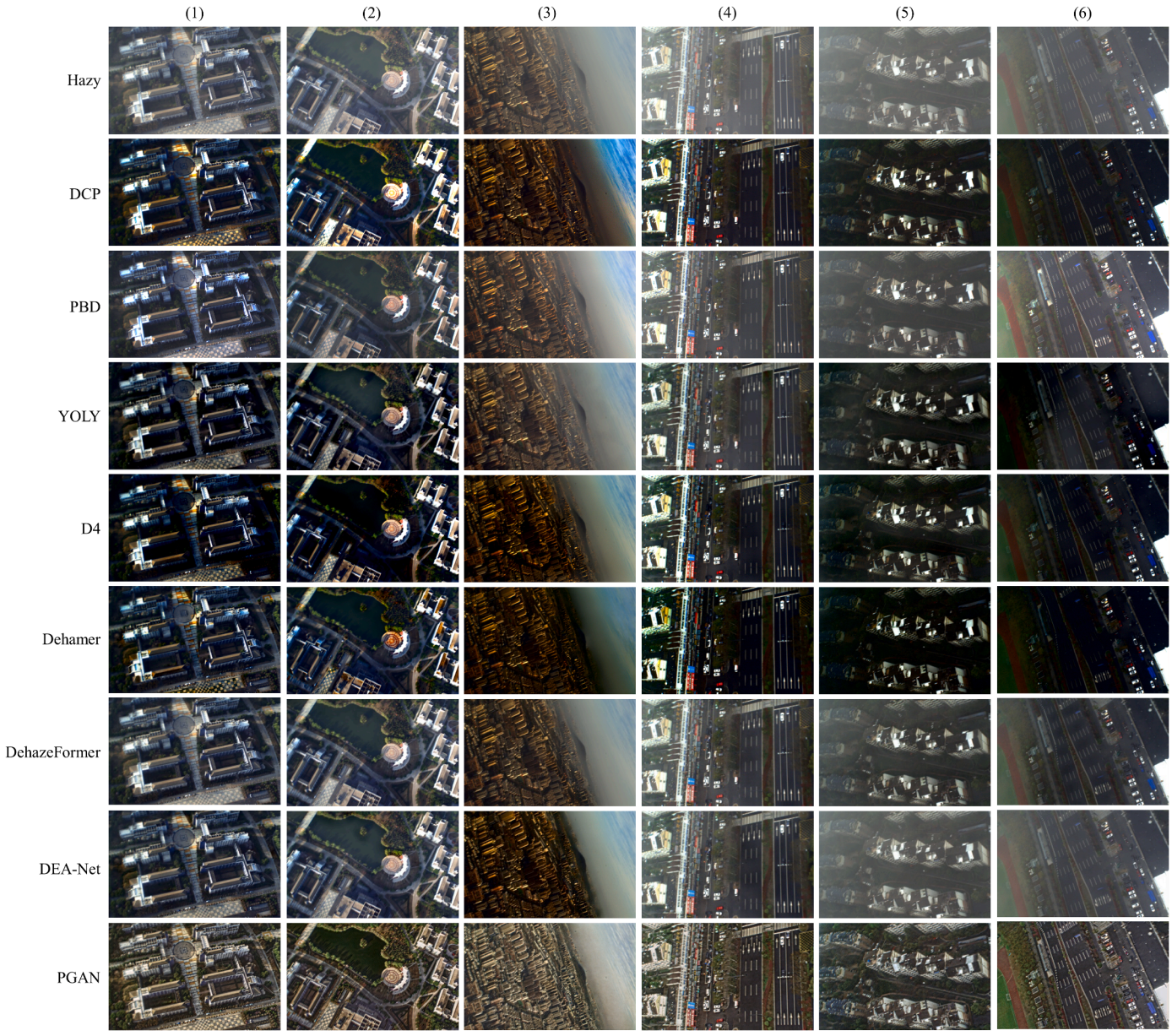


Fig. 9. Qualitative experimental results of the data in vertical direction. It can be observed that our proposed PGAN demonstrates a better dehazing effect than other polarization-based and deep-learning-based state-of-the-art methods. Data-driven deep learning-based method lacks physical model constraints and demonstrates poorer dehazing results in the absence of vertical view training data. However, PGAN exhibits a robust dehazing capability in different views.

PGAN. The experimental results of two groups are shown in Figs. 8 and Fig. 9, respectively.

In the horizontal experimental design, the haze concentration in the selected experimental scenarios (1) to (6) showed an increasing trend, in order to facilitate the testing of the performance of various algorithms under different haze concentrations. In the selection of scenes, we deliberately arranged many areas that are difficult to handle in the algorithm, such as water surfaces, bright white buildings, white trucks, etc., and some of these scenes were specifically shot under low light conditions. This test data organization is significantly different from the overly single data organization form in existing work. We hope that through this experimental data organization form, the performance of various algorithms can be more comprehensively reflected.

Horizontal direction comparison. The results of DCP and PBD are very different in thin haze scenes or containing white objects, such as scenes 1 and 2. This can be attribute to the reason that DCP is prone to estimating a higher airlight radiance from the infinite distance. In the dense haze condition, such as scenes 5 and 6, their results are similar due to the strong haze coverage.

It is obvious that PGAN has achieved excellent results in all horizontal direction images. Even in dark light and dense haze conditions, such as scenes 4 to 6, the improvement of visibility is very effective. This is made possible by the introduction of polarization calculation module and frequency information, which constrains the model to recover the hazy image as clearly and reasonably as possible under the atmospheric scattering model. Other image noises are inevitably generated in the experimental results of each method. We constantly adjust the polarization degree correction factor of PBD, but the appearance of local color shift cannot be avoided. Deep-learning based methods can better restore image details in thin haze scenes, but YOLY and D4 are worse than the advanced DEA-Net and DehazeFormer in brightness recovery. In heavy and dense haze scenes, almost all deep-learning based methods demonstrate a poor dehazing effect. Dehamer got heavier color shift in scene 5, even a seriously mistake result in scene 6. In addition to

restoring clearer dehazed images under dense haze condition, PGAN can also enhance the visibility of targets under low light conditions. Although it seems to have some haze residue and saturation reduce, it does not affect the first visual feeling of human eyes.

Vertical direction comparison. The visibility of hazy images taken in the vertical direction is related to the height of the UAV. Most of the methods such as PBD, DeahzeFormer and PGAN achieve good results from scenes 1 to 3, because these hazy images are collected in sunny days. But in scenes 4 to 6, only PBD and PGAN can keep high brightness, because there are many areas with very weak light illumination or objects with weak reflective effects in these scenes. For example, the trees under the tall buildings are shielded from most of the light in scene 5, and the asphalt road with low light reflection in scene 4 and 6. Compared with all other methods, the experimental results of PGAN are clearer, and the target recovery of the low light scene is excellent, especially the details of the blue dome of scene 5 is well restored. The trees and paths are also clearly visible. These results demonstrate that our method maintains excellent dehazing capability in a variety of scenarios.

4.3. Quantitative experimental results

We employ the widely recognized no-reference image quality assessment to quantitatively evaluate the dehazing results of our method in comparison to other dehazing methods. It is important to note that, unlike dehazing algorithms that can utilize synthetic data, polarization-based dehazing algorithms cannot directly apply reference image quality assessments for quantitative evaluation of dehazed images. Although some works have suggested collecting haze-free data at different times in the same scene or using experimental results of a certain algorithm as ground-truth, the ground-truth obtained through this method is not strict, so these practices have not yet been widely recognized. Our selected no-reference image quality assessments include Fog Aware Density Evaluator (FADE) [26], Blind Image Quality Index (BIQI) [28], Blind Image Spatial Quality Evaluator (BRISQUE) [29] and Nature Image Quality Evaluator (NIQE) [30]. These indexes based on the natural scene statistics have different priorities, allowing them to comprehensively reflect the dehazing performance rather than focusing on visibility improvement. US represents the percentage of votes in user study. The mean score of each dehazing algorithm in different groups is listed in Table 1 and Table 2. The best scores are marked in bold font.

In the horizontal direction, PGAN has achieved significant advantages in all evaluation metrics. Physical model-based methods such as PBD and DCP achieved better FADE than deep learning-based methods, indicating higher visibility in dehazing images. This proves that physical constraints are the key factor in dehazing, therefore our PGAN can achieve optimal FADE after introducing polarization information constraint. Similarly, deep learning-based methods perform poorly on BRISQUE, with a score of 16.19 indicating that our method achieves minimal image distortion and achieves the best human visual effects.

Table 2
Quantitative experimental results of the data in vertical direction. No-reference image quality assessments are calculated.

Methods	FADE↓	BIQI↓	BRISQUE↓	NIQE↓	US↑
Hazy	1.98	75.63	34.88	7.00	-
DCP [2]	0.58	53.44	26.62	5.84	0.12
PBD [17]	1.18	52.14	29.37	5.71	0.10
YOLY [32]	0.47	**42.02**	38.16	5.26	0.23
D4 [24]	0.92	61.20	29.56	6.04	0.19
Dehamer [11]	0.79	42.27	42.17	5.93	0.07
DehazeFormer [33]	1.04	73.60	31.59	6.54	0.08
SGDRL [34]	1.12	72.46	30.32	6.17	0.10
DEA-Net [35]	1.18	73.23	31.07	6.56	0.07
PGAN (Ours)	**0.37**	44.24	**22.01**	**5.22**	**0.52**

The optimal NIQE shows that our method can preserve the natural attributes of the image (e.g. contrast, clarity) during the dehazing process.

While in the vertical direction, PGAN has attained the best FADE, BRISQUE and NIQE. PGAN only achieve the second best BIQI, demonstrating that it contains some noise when generating dehazed images. This is an inevitable problem for generative-based method, such as the high BIQI value of D4. YOLY achieves the best BIQI through a layer disentanglement network that does not require training. The US scores reflect that the PGAN is the most recognized among user groups. These results demonstrate that our method can achieve robust dehazing performance in natural scenes with multiple perspectives and uneven haze.

4.4. Ablation study

The ablation experiments are performed to verify the effectiveness of our generator based on polarization and discriminator based on frequency distribution. For the generator, we remove the airlight Stokes calculation module in PGAN-A, and replace the input of PGAN-A with the original hazy image to estimate the airlight. This method is denoted as w/o SP. For the discriminator, we remove the frequency decomposition module, and only use the original and synthetic hazy images as input samples. Meanwhile, the pseudo airlight coefficient supervision loss is removed. This method is denoted as w/o FS.

As shown in Fig. 10, in the absence of the Stokes calculation module, the calculation accuracy of the airlight decreases after losing the polarization information. As a result, the dehazing performance of w/o SP is very limited. In fact, this is similar to most dehazing methods based on generative adversarial networks. If only the frequency decomposition module of our discriminator is removed, although the visibility of the scene can be improved overall, the experimental results of w/o FS are obviously less restored than PGAN in detail. It can be observed that details of trees and buildings marked by red boxes are blurred, and these details are restored in the experimental results of PGAN.

Quantitative analysis of experimental results for the three protocols on published datasets is shown in Table 3. It is not difficult to find that

Table 1
Quantitative experimental results of the data in horizontal direction. No-reference image quality assessments are calculated.

Type	Methods	Publication	FADE↓	BIQI↓	BRISQUE↓	NIQE↓	US↑
Original image	Hazy	-	6.73	85.16	45.81	9.78	-
Prior	DCP [2]	TPAMI'10	0.67	58.86	18.74	5.4	0.15
	PBD [17]	AO'03	0.65	59.9	16.74	4.81	0.27
	CAP [4]	TIP'15	3.48	78.63	20.01	7.71	0.12
Deep-Learning	AOD-Net [6]	ICCV'17	3.06	66.94	24.49	9.16	0.18
	FD-GAN [9]	AAAI'20	1.66	71.41	26.06	6.07	0.11
	YOLY [32]	IJCV'21	2.08	71.55	33.1	7.47	0.25
	D4 [24]	CVPR'22	1.54	75.18	33.68	7.58	0.09
	Dehamer [11]	CVPR'22	1.5	53.94	30.62	6.54	0.2
	DehazeFormer [33]	TIP'23	1.89	82.98	40.4	9.15	0.17
	SGDRL [34]	NN'24	2.76	79.77	34.58	8.36	0.12
	DEA-Net [35]	TIP'24	2.52	83.31	37.46	9.28	0.08
	PGAN	Ours	**0.43**	**50.01**	**16.19**	**4.73**	**0.43**

Fig. 10. Qualitative results of the ablation experiment. "w/o SP" means removing the Stokes calculation module of PGAN-A and replacing the input with the original hazy image. "w/o FS" represents for removing the frequency decomposition module of discriminator and the pseudo airlight coefficient supervision loss.

Table 3
Quantitative results of the ablation experiment. No-reference image quality assessments are calculated.

	FADE↓	BIQI↓	BRISQUE↓	NIQE↓	US↑
Hazy	6.68	90.49	35.87	10.77	-
w/o SP	4.59	64.13	23.30	8.62	0.12
w/o FS	1.78	55.43	14.54	4.77	0.41
PGAN	**0.59**	**44.65**	**12.45**	**3.95**	**0.47**

the introduction of polarization characteristics is the greatest improvement on the fog removal performance of the method while the frequency domain decomposition module can further optimize the overall performance of the network. This is consistent with our original intention of adopting the frequency decomposition module, and the accurate airlight helps to restore the image clearer and reduce the distortion.

5. Conclusion

In this work, to address the limitations of real-world labels and improve the robustness of deep learning-based dehazing methods, we propose a self-supervised polarization image dehazing framework using frequency domain generative adversarial networks, referred to as PGAN. The polarization-based generator reconstructs the synthetic hazy image using Stokes parameters and the dehazed image produced by the densely connected encoder-decoder, enabling self-supervised learning. The frequency distribution-based discriminator is more effective due to the frequency separated samples and pseudo airlight coefficient supervision loss. Benefit from incorporating the polarization and frequency information, PGAN can more accurately estimate the airlight, producing more detailed dehazed images even without real-word clear images. In experiments, PGAN has achieved advanced performance in real-world datasets, and the ablation experiments have verified the effectiveness of each module.

However, our method has limitations. Despite achieving good detail dehazing effects, certain color deviations may occur in some scenes due to the lack of supervision from ground truth.Further optimization is required to recover the image more naturally. Moreover, the proposed method lacks dynamic processing capabilities, such as video dehazing. In the future, we will delve into research to improve the dehazing ability of the model in complex scenes with varying depths.

CRediT authorship contribution statement

Rui Sun: Writing – review & editing, Methodology, Investigation, Funding acquisition, Conceptualization. **Long Chen:** Writing – original draft, Validation, Methodology, Investigation, Conceptualization. **Tanbin Liao:** Writing – review & editing, Investigation, Conceptualization. **Zhiguo Fan:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 62302142 and Grant 61876057; in part by the China Postdoctoral Science Foundation under Grant 2022M720981; in part by the Anhui Province Natural Science Foundation under Grant No 2208085MF158; and in part by the Key Research Plan of Anhui Province - Strengthening Police with Science and Technology under Grant 202004d07020012.

Data availability

Data will be made available on request.

References

[1] X. Guo, Y. Yang, C. Wang, J. Ma, Image dehazing via enhancement, restoration, and fusion: a survey, Inf. Fusion 86 (2022) 146–170.
[2] K. He, J. Sun, X. Tang, Single image haze removal using dark channel prior, IEEE Trans. Pattern Anal. Mach. Intell. 33 (2011) 2341–2353.
[3] G. Meng, Y. Wang, J. Duan, S. Xiang, C. Pan, Efficient image dehazing with boundary constraint and contextual regularization, in: in IEEE International Conference on Computer Vision (ICCV), 2013, pp. 617–624.
[4] Q. Zhu, J. Mai, L. Shao, A fast single image haze removal algorithm using color attenuation prior, IEEE Trans. Image Process. 24 (2015) 3522–3533.
[5] Wenqi Ren, et al., Single image dehazing via multi-scale convolutional Neural networks, in: European Conference on Computer Vision, 2016.
[6] B. Li, X. Peng, Z. Wang, J. Xu, F. Dan, AOD-net: all-in-one dehazing network, in: in IEEE International Conference on Computer Vision (ICCV), 2017, pp. 4770–4778.
[7] B. Li, Y. Gou, J.Z. Liu, H. Zhu, J.T. Zhou, X. Peng, Zero-shot image dehazing, in: in IEEE Transactions on Image Processing 29, 2020, pp. 8457–8466, https://doi.org/10.1109/TIP.2020.3016134.
[8] S. Zhang, X. Zhang, S. Wan, W. Ren, L. Zhao, L. Shen, Generative adversarial and self-supervised dehazing network, IEEE Trans. Ind. Informat. 20 (3) (March 2024) 4187–4197.
[9] Y. Dong, Y. Liu, H. Zhang, S. Chen, Y. Qiao, FD-GAN: generative adversarial networks with fusion-discriminator for single image dehazing, in: in AAAI Conference on Artificial Intelligence (AAAI), 2020, pp. 10729–10736.
[10] H. Dong, et al., Multi-scale boosted dehazing network with dense feature fusion, in: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 2154–2164.
[11] C. Guo, Q. Yan, S. Anwar, R. Cong, W. Ren, C. Li, Image dehazing transformer with transmission-aware 3D position embedding, in: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 2022, pp. 5802–5810.
[12] R.-Q. Wu, Z.-P. Duan, C.-L. Guo, Z. Chai, C. Li, RIDCP: revitalizing real image dehazing via high-quality codebook priors, in: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 2023, pp. 22282–22291.
[13] S.Li Liang, D. Cheng, W. Wang, D. Li, J. Liang, Image dehazing via self-supervised depth guidance, Pattern Recognit. 158 (2025) 111051.
[14] T. Jia, J. Li, L. Zhuo, G. Li, Effective Meta-attention dehazing networks for vision-based Outdoor Industrial systems, IEEE Trans. Ind. Informat. 18 (3) (March 2022) 1511–1520.

[15] Y. Cui, Q. Wang, C. Li, W. Ren, A. Knoll, EENet: an effective and efficient network for single image dehazing, Pattern Recognit. 158 (2025) 111074.
[16] N. Jiang, K. Hu, T. Zhang, W. Chen, Y. Xu, T. Zhao, Deep hybrid model for single image dehazing and detail refinement, Pattern Recognit. 136 (2023) 109227.
[17] Y.Y. Schechner, S.G. Narasimhan, S.K. Nayar, Polarization-based vision through haze, Appl. Opt. 42 (2003) 511–525.
[18] Y. Qu, Z. Zou, Non-sky polarization-based dehazing algorithm for non-specular objects using polarization difference and global scene feature, Opt. Express 25 (2017) 25004–25022.
[19] F. Huang, C. Ke, X. Wu, S. Wang, J. Wu, X. Wang, Polarization dehazing method based on spatial frequency division and fusion for a far-field and dense hazy image, Appl. Opt. 60 (2021) 9319–9332.
[20] Rui Sun, Tanbin Liao, Zhiguo Fan, Xudong Zhang, Changxiang Wang, Polarization dehazing method based on separating and iterative optimizing airlight from the frequency domain for different concentrations of haze, Appl. Opt. 61 (2022) 10362–10373.
[21] A. Dudhane, H.S. Aulakh, S. Murala, RI-GAN: an end-to-end network for single image haze removal, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Long Beach, CA, USA, 2019, pp. 2014–2023.
[22] A. Dudhane, S. Murala, CDNet: single image de-hazing using unpaired adversarial training, in: 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 2019, pp. 1147–1155.
[23] W. Liu, X. Hou, J. Duan, G. Qiu, End-to-End single image fog removal using enhanced cycle consistent adversarial networks, in: in IEEE Transactions on Image Processing 29, 2020, pp. 7819–7833.
[24] Y. Yang, C. Wang, R. Liu, L. Zhang, X. Guo, D. Tao, Self-augmented unpaired image dehazing via density and depth decomposition, in: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 2022, pp. 2027–2036.
[25] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (2004) 600–612.
[26] L.K. Choi, J. You, A.C. Bovik, Referenceless prediction of perceptual fog density and perceptual image defogging, IEEE Trans. Image Process. 24 (2015) 3888–3901.
[27] A.L. Da Cunha, J. Zhou, M.N. Do, The nonsubsampled contourlet transform: theory, design, and applications, IEEE Trans. Image Process. 15 (2006) 3089–3101.
[28] A.K. Moorthy, A.C. Bovik, A two-step framework for constructing blind image quality indices, IEEE Signal Process. Lett. 17 (2010) 513–516.
[29] A. Mittal, A.K. Moorthy, A.C. Bovik, No-reference image quality assessment in the spatial domain, IEEE Trans. Image Process. 21 (2012) 4695–4708.
[30] A. Mittal, R. Soundararajan, A.C. Bovik, Making a "completely blind" image quality analyzer, IEEE Signal Process. Lett. 20 (2012) 209–212.
[31] Y. Wang, et al., UCL-dehaze: toward real-world image dehazing via unsupervised contrastive learning, in IEEE Trans Image Process 33 (2024) 1361–1374, https://doi.org/10.1109/TIP.2024.3362153.
[32] B. Li, Y. Gou, S. Gu, J.Z. Liu, J.T. Zhou, X. Peng, You only look yourself: unsupervised and untrained single image dehazing neural network, Int. J. Comput. Vis. 129 (2021) 1754–1767.
[33] Y. Song, Z. He, H. Qian, X. Du, Vision transformers for single image dehazing, in IEEE Trans. Image Process. 32 (2023) 1927–1941.
[34] T. Jia, J. Li, L. Zhuo, J. Zhang, Self-guided disentangled representation learning for single image dehazing, Neural Netw. 172 (2024) 106107.
[35] Z. Chen, Z. He, Z.-M. Lu, DEA-net: single image dehazing based on detail-enhanced convolution and content-guided attention, in IEEE Trans. Image Process. 33 (2024) 1002–1015.
[36] Cheng Wang, Zhiguo Fan, Jin Haihong, Xianqiu Wang, Dou Hua, Design and optimization analysis of imaging system of polarized skylight pattern of full polarization, Acta Phys. Sin. 70 (2021), 104201-1-104201-9.
[37] S. Liang, T. Gao, T. Chen, P. Cheng, A remote sensing image dehazing method based on heterogeneous priors, IEEE Trans. Geosci. Remote Sens. 62 (2024) 1–13.
[38] Z. Lyu, Y. Chen, Y. Hou, MCPNet: multi-space color correction and features prior fusion for single-image dehazing in non-homogeneous haze scenarios, Pattern Recognit. 150 (2024) 110290.
[39] H. Sun, B. Li, Z. Dan, et al., Multi-level feature interaction and efficient non-local information enhanced channel attention for image dehazing, Neural Netw. 163 (2023) 10–27.
[40] Z. Qiu, T. Gong, Z. Liang, et al., Perception-oriented UAV image dehazing based on super-pixel scene prior, IEEE Trans. Geosci. Remote Sens. 62 (2024) 1–19.
[41] H. Sun, Y. Wen, H. Feng, et al., Unsupervised bidirectional contrastive reconstruction and adaptive fine-grained channel attention networks for image dehazing, Neural Netw. 176 (2024) 106314.
[42] R. Sun, T. Liao, Z. Fan, X. Zhang, and C. Wang, "Polarization dehazing datasets for different concentrations of haze," in figshare, 2022. https://doi.org/10.6084/m9.figshare.21644927.

Sui Sun received the B.S. degree from Central South University of China in 1998, the M.S. degree from Harbin Engineering University of China in 2000, the Ph.D. degree from Huazhong university of Science and Technology of China in 2003. He was a visiting scholar in Computer Science department, University of Missouri-Columbia, USA from 2010 to 2011. He is currently a professor in the School of Computer and Information of Hefei University of Technology, China. His research interests include object recognition and tracking, computer vision, and machine learning.

Long Chen received the B.S. degree from Hefei University of Technology of China in 2022. He is currently pursuing M.S. degree in Hefei University of Technology. His research interests include machine learning and computer vision, especially cross modal person re-identification and corruption robustness.

Tanbin Liao received the B.S. degree from Hefei University of Technology of China in 2021. He is currently pursuing M.S. degree in Hefei University of Technology. His research interests include machine learning and computer vision, especially polarization image dehazing.

Zhiguo Fan was born in Anhui, China, in 1978. He received the M.S. and Ph.D. degrees from the Hefei University of Technology in 2007 and 2011, respectively. He is a Professor with the School of Computer and Information, Hefei University of Technology. He involved in bionic polarized light navigation, polarization optical detection, the research of intelligent information processing theory, and development of the related application technology. In recent years, he has been authorized 17 invention patents, published >40 academic articles and one academic monograph [Bionic Polarized Light Navigation Method (Science Press), 2014].