

HFIP: High-Frequency Information Prompt Network for Single Image Dehazing

Jufeng Li*, Yuxin Feng*, Mouguang Lin*, Zhuo Su*[†]

*School of Computer Science and Engineering, Sun Yat-sen University
Guangzhou, China

[†]Corresponding author: suzhuo3@mail.sysu.edu.cn

Abstract—Current dehazing methods perform well on synthetic haze datasets but often struggle with complex real-world scenarios. To enhance the model’s generalization ability, we incorporate Fourier transform and Gaussian filter to capture frequency domain information. Based on this, we propose a high-frequency information prompt network to more effectively locate haze and restore texture details. Additionally, to address the issue of overly dark dehazing results commonly observed in networks trained on synthetic datasets, we introduce a brightness loss function based on threshold division, which effectively enhances the brightness of the output images. Through extensive experiments, our method demonstrates not only excellent dehazing performance in challenging and complex scenarios such as daytime fog and sandstorm, but also significant potential value for downstream applications.

Index Terms—Image Dehazing, Fourier Transform, Gaussian Filter, High-Frequency Information

I. INTRODUCTION

Image dehazing aims to improve the visual quality of images taken under hazy weather conditions. While existing single image dehazing methods based on deep learning [1], [2], [3] have achieved impressive results in relatively simple and controlled scene, they often fall short in handling the complexities of real-world hazy scenarios. A key limitation of these methods is their inability to accurately locate haze, resulting in residual haze and significant loss of texture details, which diminishes image clarity and leads to structural distortion. To address these challenges, we propose a novel image dehazing network based on high-frequency information prompt. This approach enhances the network’s capability to accurately locate haze and restore image texture details, thereby improving overall dehazing performance.

The motivation behind our proposed method stems from our observation that high-frequency information in an image retains more texture details and contains almost no haze, whereas a significant amount of haze is present in low-frequency information.

As illustrated in Figure 1, we utilize the fast Fourier transform to map both hazy and clear images into the frequency domain. Then we apply Gaussian filter (GF) to decompose these images into high-frequency and low-frequency information, followed by an inverse fast Fourier transform to reconstruct the high-frequency images (High) and low-frequency images (Low), respectively. The results show that the high-frequency image contains more texture details with almost no visible

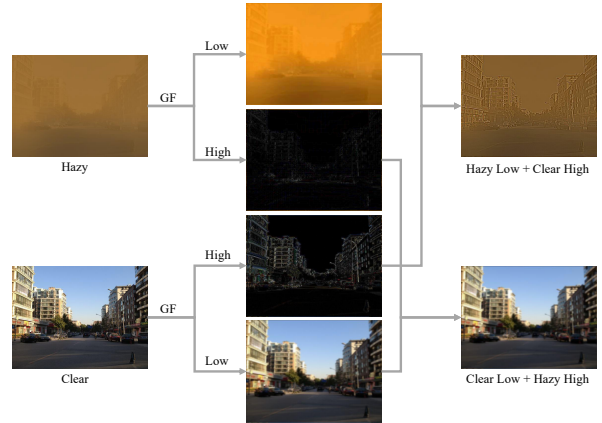


Fig. 1. By swapping the high-frequency and low-frequency information between a clear image and a hazy image, we observe that the high-frequency information retains more texture details and less haze.

haze, while the low-frequency image captures significant background information along with yellow haze. Additionally, we superimpose the low-frequency signal of the hazy image with the high-frequency signal of the clear image (denoted as **HLCH**), and the high-frequency signal of the hazy image with the low-frequency signal of the clear image (denoted as **CLHH**), and then perform an inverse fast Fourier transform on these combined signals. It is observed that **HLCH** retains considerable yellow haze, with enhanced texture details and a visual similarity to the original hazy image. In contrast, **CLHH** shows almost no yellow haze, with blurred texture details, making it visually closer to the clear image.

To further validate the observation, we employ the haze-related metric FRFSIM [4] for a statistical analysis of 500 pairs of training data from the OTS dataset [5]. This metric measures the similarity between the regenerated image and the input hazy image, with higher scores indicating more haze in the regenerated image. The average FRFSIM score for **HLCH** compared to the input hazy image is 0.6396, whereas the average FRFSIM score for **CLHH** is only 0.4242. These results show that Gaussian filter can effectively separate texture details from haze in hazy images, providing valuable guidance for the network to produce clearer and more natural dehazing results.

Additionally, datasets synthesized using the atmospheric

light scattering model (ASM) do not account for light absorption, leading to hazy images that are consistently brighter than their clear counterparts. As a result, networks trained on these synthetic datasets often produce darker result when applied to real hazy scenarios. To address this issue, we design a brightness loss based on threshold division into the loss function. Experimental results show that this adjustment effectively reduces the network's tendency to generate darker outputs, thereby enhancing the overall brightness of the images.

Our main contributions are as follows:

- To enhance the network's ability to locate haze and restore image texture details, we leverage frequency domain information obtained through Fourier transform and Gaussian filter to develop a dehazing network that utilizes high-frequency information prompt.
- We introduce a novel brightness loss function based on threshold division to enhance the overall brightness of the images, improving visual clarity and detail.
- We perform extensive experiments on real-world haze dataset, which thoroughly validate the effectiveness and robustness of our proposed method in handling haze in real-world scenarios.

II. RELATED WORK

In this section, we review and summarize existing works on single-image dehazing from two perspectives: dataset construction and training strategies.

A. Dataset Construction

The development of realistic synthetic datasets is crucial for enhancing the real-world dehazing performance of networks. Early efforts primarily involved using the atmospheric scattering model (ASM) to generate synthetic datasets, such as the RESIDE dataset synthesized by Li et al. [5], which remains one of the largest haze datasets. To better simulate real-world hazy conditions, Wu et al. [6] designed a haze training dataset incorporating multiple degradation factors. Additionally, Ancuti et al. [7] developed the indoor dataset I-Haze and the outdoor dataset O-Haze [8] using a haze generator to replicate real-world hazy scenarios. These datasets have significantly advanced dehazing research.

B. Training Strategies

From the perspective of training strategies, deep learning-based dehazing methods can be categorized into paired and unpaired data-based approaches.

Paired Data-Based Training: This supervised approach involves training an end-to-end convolutional neural network (CNN) with paired hazy and clear images. For instance, Zheng et al. [3] developed an end-to-end contrast-regularized physical perception dehazing network utilizing contrastive learning. Feng et al. [9] introduced a network specifically designed for haze localization and removal. Additionally, incorporating image prior information into the network has been shown to enhance dehazing performance. Wang et al. [2] proposed a self-prompted dehazing network leveraging image depth

consistency, and Feng et al. [10] utilized two-dimensional discrete wavelet prior features.

Unpaired Data-Based Training: To enhance the model's generalization capability, training based on unpaired data has also gained significant attention in recent years. This category includes semi-supervised and unsupervised methods. Unsupervised approaches often employ generative adversarial networks (GANs). For example, Li et al. [11] proposed using conditional GANs for dehazing. Semi-supervised methods are also effective in various image processing tasks. For instance, Li et al. [12] aligned synthetic and real data in a high-dimensional space, and Cong et al. [13] proposed a semi-supervised network for nighttime dehazing.

III. METHOD

A. Network Architecture

Figure 2 and Figure 3 illustrate the overall pipeline of the proposed HFIP network and the HFII module, respectively. Next, we describe the computational process of the HFII module in detail.

For an input feature $x \in \mathbb{R}^{H \times W \times C}$ at the current layer, we first apply the fast Fourier transform to efficiently map x to the frequency domain:

$$\mathbb{F}(x)(u, v) = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} x(h, w) \exp \left(-j2\pi \left(\frac{h}{H}u + \frac{w}{W}v \right) \right). \quad (1)$$

Here, (h, w) and (u, v) denote coordinates in the spatial domain and frequency domain, respectively. The inverse fast Fourier transform is denoted as $\mathbb{F}^{-1}$. We then apply Gaussian filter in the frequency domain to extract high-frequency information (HFI) and low-frequency information (LFI):

$$\begin{aligned} HFI &= \mathbb{F}^{-1} \left(\mathbb{F}(x) \cdot \exp \left(-\frac{H^2 + W^2}{2k^2} \right) \right), \\ LFI &= \mathbb{F}^{-1} \left(\mathbb{F}(x) \cdot \left(1 - \exp \left(-\frac{H^2 + W^2}{2k^2} \right) \right) \right), \end{aligned} \quad (2)$$

where the kernel size $k = \min(H, W)/10$.

To enhance the network's ability to locate haze and repair texture details, we incorporate both spatial attention and channel attention [14] ($CA(\cdot)$) mechanisms. Specifically, we apply the spatial attention mechanism to HFI to derive a weight map that emphasizes texture details, denoted as $W_H = SA(HFI)$. Similarly, the channel attention mechanism is applied to LFI to derive a weight map that focuses on haze, denoted as $W_L = CA(LFI)$. These attention weights are then applied to the input feature x to obtain the output of FAM as follows:

$$x^* = W_L \times (W_H \times x). \quad (3)$$

Inspired by Wang et al. [2], we enhance the application of window-based multi-head self-attention (W-MSA) to the feature x^* . Specifically, we linearly superimpose the high-frequency information (HFI) onto the query (Q) to obtain a modified query Q' :

$$Q' = Q + HFI. \quad (4)$$

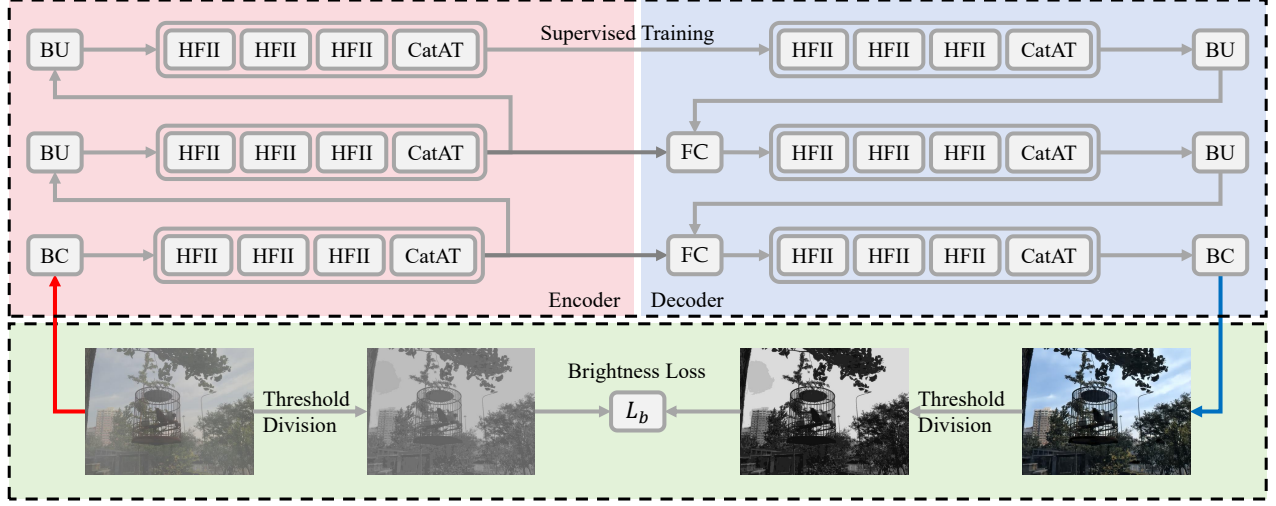


Fig. 2. The proposed network HFIP mainly consists of high-frequency information interaction modules (HFII) and includes basic convolution (BC), sampling unit (BU), and fully connected layer (FC). The CatAT module fuses outputs from three consecutive HFII modules.

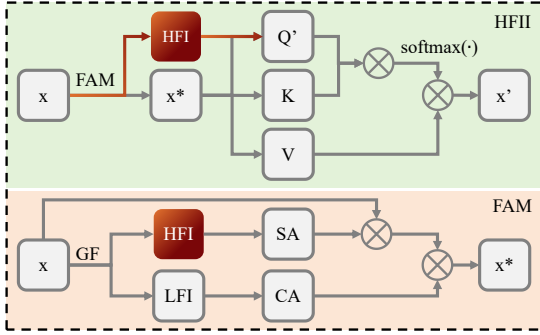


Fig. 3. The structure of the high-frequency information interaction module (HFII) and its submodule, the frequency domain attention module (FAM).

This approach enables the high-frequency information to act as a prompt, guiding the network to focus more on the texture details within the image.

B. Training loss optimization

The input hazy images and their corresponding clear labels are denoted as X and Y . Our proposed HFIP network is represented by Φ .

The network output and clear labels are aligned in both the spatial domain and frequency domain:

$$\begin{aligned} \mathcal{L}_s &= \| \Phi(X) - Y \|_1, \\ \mathcal{L}_f &= \| \mathbb{F}(\Phi(X)) - \mathbb{F}(Y) \|_1. \end{aligned} \quad (5)$$

Furthermore, to enhance visual perception, we incorporate a multi-scale structural similarity loss [15]:

$$\mathcal{L}_{ms} = 1 - \prod_{s=1}^S \left(\frac{2\mu_i\mu_j + c_1}{\mu_i^2 + \mu_j^2 + c_1} \right)^\alpha \left(\frac{2\sigma_{ij} + c_2}{\sigma_i^2 + \sigma_j^2 + c_2} \right)^\beta. \quad (6)$$

Here, S denotes different scales. μ_i and μ_j denote the mean values of $\Phi(X)$ and Y respectively. σ_i and σ_j denote the standard deviation of $\Phi(X)$ and Y respectively. σ_{ij} represents the covariance between $\Phi(X)$ and Y . α, β control the relative importance of two items, and c_1 and c_2 are constant values.

As shown in Figure 2, to enhance the brightness of the dehazing result, we approximate the image's depth map using the blue channel [16]. We numerically divide this depth map into n regions and compute the average pixel value for each corresponding region in the foggy image. For the m -th region Ω_m of the foggy image X , the average brightness is given by:

$$\psi(X_m) = \frac{1}{3M} \sum_{c=0}^2 \sum_{i,j \in \Omega_m} X_m(i, j, c), \quad (7)$$

where M is the number of pixels in Ω_m . The brightness of the corresponding region in the dehazing image $\Phi(X)$ is calculated similarly. We use the parameter γ to adjust the brightness upper limit, and then average the brightness values across all regions to obtain the overall brightness loss:

$$\mathcal{L}_b = \frac{1}{n} \sum_{m=1}^n \| \psi(\Phi(X_m)) - (\psi(X_m))^\gamma \|_1, \quad (8)$$

where $n = 10$ and $\gamma = 1.3$.

Finally, the total loss function is defined as:

$$\mathcal{L}_{total} = \mathcal{L}_s + \alpha \mathcal{L}_f + \mathcal{L}_{ms} + \mathcal{L}_b, \quad (9)$$

where $\alpha = 0.1$.

IV. EXPERIMENTAL RESULTS

A. Implementation Details

1) *Compared Methods*: We compare our method with several state-of-the-art data-driven supervised dehazing methods, including TOE [17], C2P [3], DEA [18], SGID [19], and WeatherDiff (WD) [20].



Fig. 4. Qualitative results of advanced dehazing methods on the RTTS dataset demonstrate that our method excels in dehazing scenes affected by daytime fog and sandstorms.

2) *Dataset Description*: We construct a dataset with 37,129 training pairs, covering various haze types such as daytime fog, nighttime fog, sandstorm, and colorful fog. To assess real-world performance, we evaluate different methods on the RTTS dataset [5] with 4,322 images.

3) *Evaluation Metrics*: We use five authoritative no-reference image evaluation metrics to assess the performance of different methods on real-world hazy scenarios: Fog Aware Density Evaluator (FADE) [21], Picture-based Predictor of PM2.5 Concentration (PPPC) [22], Natural Image Quality Evaluator (NIMA) [23], Image Entropy, and Blind Image Quality Metric for Enhanced Images (BIQME) [24].

B. Performance Evaluation

As shown in Figure 4, our method produces richer texture details and more noticeable dehazing effect in daytime hazy scenarios compared to other approaches. SGID removes some haze but at the cost of losing texture details, while TOE suffers from significant color distortion. In sandstorm scenarios, our method delivers the best dehazing effect, whereas TOE only partially removes haze, and other methods are powerless. Table I demonstrates that our method ranks among the top two in all five authoritative no-reference image quality indicators.

Moreover, we evaluate the practical application of different dehazing methods by applying their results to downstream task. We use YOLOv8n for object detection on the RTTS dataset. As shown in Figure 5, in hazy road scenarios, our method not only detects more vehicles, but also achieves higher confidence scores and a lower false detection rate. For instance, in the second row, other methods misidentify bus while our method correctly detects more cars. Table II presents the numerical results for object detection, where our method

TABLE I
QUANTITATIVE RESULTS OF ADVANCED DEHAZING METHODS ON THE RTTS DATASET ARE PRESENTED USING FIVE NO-REFERENCE IMAGE QUALITY METRICS. THE TOP TWO SCORES FOR EACH METRIC ARE HIGHLIGHTED IN BOLD AND UNDERLINED.

Metrics	C2P	DEA	SGID	TOE	WD	Ours
FADE↓	2.0606	1.7939	1.6441	1.1319	2.4113	1.4914
PPPC↓	188.05	183.25	178.36	<u>164.16</u>	193.27	160.18
NIMA↑	3.7139	3.8116	3.6941	<u>3.9794</u>	3.6421	4.8165
Entropy↑	7.1682	7.1911	<u>7.2199</u>	<u>7.1747</u>	7.1443	7.3150
BIQME↑	0.5305	0.5404	0.5452	<u>0.5690</u>	0.5158	0.5722

TABLE II
QUANTITATIVE RESULTS OF OBJECT DETECTION ON THE RTTS DATASET ARE PRESENTED. OUR METHOD ACHIEVES EXCELLENT METRICS ACROSS ALL EVALUATED INDICATORS.

Methods	Person	Bicycle	Car	Motorbike	Bus	mAP
Hazy	0.5977	0.3543	0.4540	0.2617	0.1940	0.3723
C2P	0.6074	0.3598	0.4598	0.2969	0.2020	0.3852
DEA	0.6106	0.3674	0.4629	0.3058	0.2030	0.3900
SGID	<u>0.6143</u>	0.3741	<u>0.4715</u>	<u>0.3071</u>	0.2086	<u>0.3951</u>
TOE	0.6113	0.3497	0.4646	0.2988	<u>0.2147</u>	0.3878
WD	0.5017	0.3049	0.3403	0.2156	0.1705	0.3066
Ours	0.6431	<u>0.3724</u>	0.5100	0.3486	0.2362	0.4221

ranks among the top two for each detection metric, with a significantly higher final mAP score.

C. Ablation Study

To evaluate the effectiveness of our method, we conduct the following ablation experiments: (1) Removing the frequency domain attention module (FAM). (2) Removing the concatenation attention module (CatAT). (3) Removing the high-frequency information prompt (HFI). (4) Removing the brightness loss term (L_b).

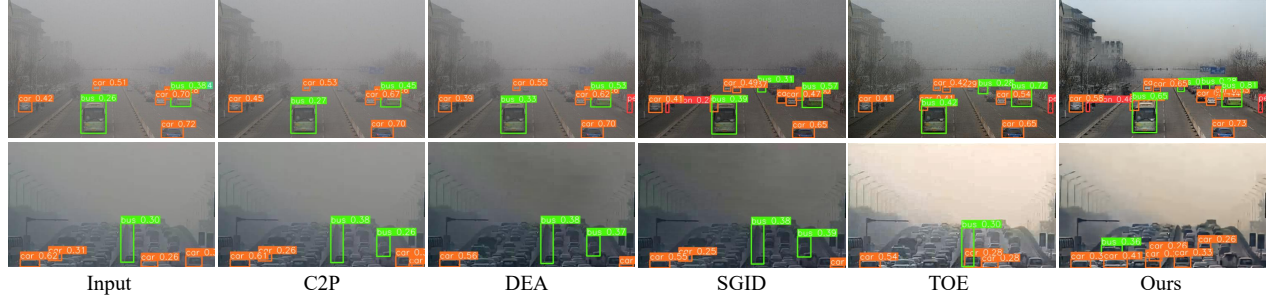


Fig. 5. Qualitative results of vehicle detection using YOLOv8n in real foggy road scenes demonstrate that our method significantly improves detection performance.

TABLE III
THE IMPACT OF VARIOUS COMPONENTS ON THE PERFORMANCE OF THE NETWORK MODEL.

Model	PSNR $\uparrow$	SSIM $\uparrow$	Params	FLOPs
w/o FAM	18.39	0.88	1.31M	15.01G
w/o CatAT	22.18	0.87	1.28M	15.06G
w/o HFI	19.71	0.89	1.33M	15.11G
w/o L_b	23.44	0.91	1.33M	15.11G
Full model	25.38	0.93	1.33M	15.11G

We use Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) to evaluate the effectiveness of different modules on the synthetic SOTS dataset [5] with 500 images. As shown in Table III, removing the frequency domain attention module (FAM) or the high-frequency information prompt significantly degrades dehazing performance. Moreover, the FAM has a very small parameter count, and the high-frequency information prompt does not require any additional parameters, demonstrating the efficiency and effectiveness of our module design.

To thoroughly demonstrate the effectiveness of brightness loss, we conducted a visual evaluation on the RTTS and URHI dataset [5]. As shown in Figure 6, introducing brightness loss not only improves the dehazing effect but also enhances the overall brightness of the image.

V. CONCLUSION

To address the dehazing problem in real-world scenarios, we use Fourier transform and Gaussian filter to extract haze and texture information from images, guiding the network to produce clearer dehazing results. Additionally, we design a threshold-based brightness loss function to enhance image brightness. Extensive experiments on real hazy scenarios demonstrate the effectiveness and robustness of our method.

ACKNOWLEDGMENT

This work was supported in part by the Guangxi Science and Technology Plan Project under Grant No. GuikeAD23026034, and in part by the Guangdong Basic and Applied Basic Research Foundation under Grant No. 2024A1515011563.



Fig. 6. The impact of brightness loss on dehazing results.

REFERENCES

- [1] Y. Feng, X. Meng, F. Zhou, W. Lin, and Z. Su, "Real-world non-homogeneous haze removal by sliding self-attention wavelet network," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 33, no. 10, pp. 5470–5485, 2023.
- [2] C. Wang, J. Pan, W. Lin, J. Dong, W. Wang, and X.-M. Wu, "Selfpromer: Self-prompt dehazing transformers with depth-consistency," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 6, pp. 5327–5335, Mar. 2024. [Online]. Available: <https://ojs.aaai.org/index.php/AAAI/article/view/28340>
- [3] Y. Zheng, J. Zhan, S. He, J. Dong, and Y. Du, "Curricular contrastive regularization for physics-aware single image dehazing," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2023, pp. 5785–5794.
- [4] W. Liu, F. Zhou, T. Lu, J. Duan, and G. Qiu, "Image defogging quality assessment: Real-world database and method," *IEEE Transactions on Image Processing*, vol. 30, pp. 176–190, 2021.
- [5] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang, "Benchmarking single-image dehazing and beyond," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 492–505, 2019.
- [6] R.-Q. Wu, Z.-P. Duan, C.-L. Guo, Z. Chai, and C. Li, "Ridcep: Revitalizing real image dehazing via high-quality codebook priors," in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023, pp. 22 282–22 291.
- [7] C. Ancuti, C. O. Ancuti, R. Timofte, and C. De Vleeschouwer, "I-haze: A dehazing benchmark with real hazy and haze-free indoor images," in *Advanced Concepts for Intelligent Vision Systems*, J. Blanc-Talon, D. Helbert, W. Philips, D. Popescu, and P. Scheunders, Eds. Cham: Springer International Publishing, 2018, pp. 620–631.

- [8] C. O. Ancuti, C. Ancuti, R. Timofte, and C. De Vleeschouwer, "O-haze: A dehazing benchmark with real hazy and haze-free outdoor images," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2018, pp. 867–8678.
- [9] Y. Feng, L. Ma, X. Meng, F. Zhou, R. Liu, and Z. Su, "Advancing real-world image dehazing: Perspective, modules, and training," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–18, 2024.
- [10] Y. Feng, Z. Su, L. Ma, X. Li, R. Liu, and F. Zhou, "Bridging the gap between haze scenarios: A unified image dehazing model," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1–1, 2024.
- [11] R. Li, J. Pan, Z. Li, and J. Tang, "Single image dehazing via conditional generative adversarial network," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 8202–8211.
- [12] X. Li, Y. Feng, F. Zhou, Y. Liang, and Z. Su, "Revitalizing real image deraining via a generic paradigm towards multiple rainy patterns," in *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*, K. Larson, Ed. International Joint Conferences on Artificial Intelligence Organization, 8 2024, pp. 1029–1037, main Track. [Online]. Available: <https://doi.org/10.24963/ijcai.2024/114>
- [13] X. Cong, J. Gui, J. Zhang, J. Hou, and H. Shen, "A semi-supervised nighttime dehazing baseline with spatial-frequency aware and realistic brightness constraint," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2024, pp. 2631–2640.
- [14] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "Cbam: Convolutional block attention module," in *Proceedings of the European Conference on Computer Vision (ECCV)*, September 2018.
- [15] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [16] M. Ju, C. Ding, C. A. Guo, W. Ren, and D. Tao, "Idrlp: Image dehazing using region line prior," *IEEE Transactions on Image Processing*, vol. 30, pp. 9043–9057, 2021.
- [17] Y. Gao, W. Xu, and Y. Lu, "Let you see in haze and sandstorm: Two-in-one low-visibility enhancement network," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–12, 2023.
- [18] Z. Chen, Z. He, and Z.-M. Lu, "Dea-net: Single image dehazing based on detail-enhanced convolution and content-guided attention," *IEEE Transactions on Image Processing*, vol. 33, pp. 1002–1015, 2024.
- [19] H. Bai, J. Pan, X. Xiang, and J. Tang, "Self-guided image dehazing using progressive feature fusion," *IEEE Transactions on Image Processing*, vol. 31, pp. 1217–1229, 2022.
- [20] O. Özdenizci and R. Legenstein, "Restoring vision in adverse weather conditions with patch-based denoising diffusion models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 8, pp. 10 346–10 357, 2023.
- [21] L. K. Choi, J. You, and A. C. Bovik, "Referenceless prediction of perceptual fog density and perceptual image defogging," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3888–3901, 2015.
- [22] K. Gu, J. Qiao, and X. Li, "Highly efficient picture-based prediction of pm2.5 concentration," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 4, pp. 3176–3184, 2019.
- [23] H. Talebi and P. Milanfar, "Nima: Neural image assessment," *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 3998–4011, 2018.
- [24] K. Gu, D. Tao, J.-F. Qiao, and W. Lin, "Learning a no-reference quality assessment model of enhanced images with big data," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 4, pp. 1301–1313, 2018.