PDD-GAN: Prior-based GAN Network with Decoupling Ability for Single Image Dehazing

Xiaoxuan Chai*
Tsinghua Shenzhen International Graduate School
Shenzhen, China
cxx20@mails.tsinghua.edu.cn

Junchi Zhou*
Tsinghua Shenzhen International Graduate School
Shenzhen, China
zhoujc20@mails.tsinghua.edu.cn

Hang Zhou
PAII Inc.
Palo Alto, California, USA
joeyzhou1984@gmail.com

Juihsin Lai†
PAII Inc.
Palo Alto, California, USA
juihsin.lai@gmail.com

ABSTRACT

Single image dehazing is a challenging vision problem aiming to provide clear images for downstream computer vision applications (e.g., semantic segmentation, object detection, and super resolution). Most existing methods leverage the physical scattering model or convolutional neural networks (CNNs) for haze removal, which however ignore the complementary advantages between each other. Especially lacking marginal and visual prior instructions, CNN-based methods still have gaps in details and color recovery. To solve these, we propose a Prior-based with Decoupling ability Dehazing GAN Network (PDD-GAN), which is based on PeleetNet and attached with an attention module (CBAM). The prior-based decoupling approach consists of two parts: high and low frequency filtering and HSV contrastive loss. We process the image via a band-stop filter and add it as the fourth channel of data ($RGBF_{HL}$) to decouple the hazy image at the structural level. Besides, a novel prior loss with contrastive regularization is proposed at the visual level. Sufficient experiments are carried out to demonstrate that PDD-GAN outperforms state-of-the-art methods by up to 0.86db in PSNR. In particular, extensive experiments indicate that $RGBF_{HL}$ increases by 0.99db compared with the original three-channel data (RGB) and the extra HSV prior loss escalates by 2.0db. Above all, our PDD-GAN indeed has the decoupling ability and improves the dehazing results.

CCS CONCEPTS

• **Computing methodologies** → **Computer vision problems**; **Computer vision representations**; **Computer vision tasks**.

*Both authors contributed equally to this research.
†Corresponding author.

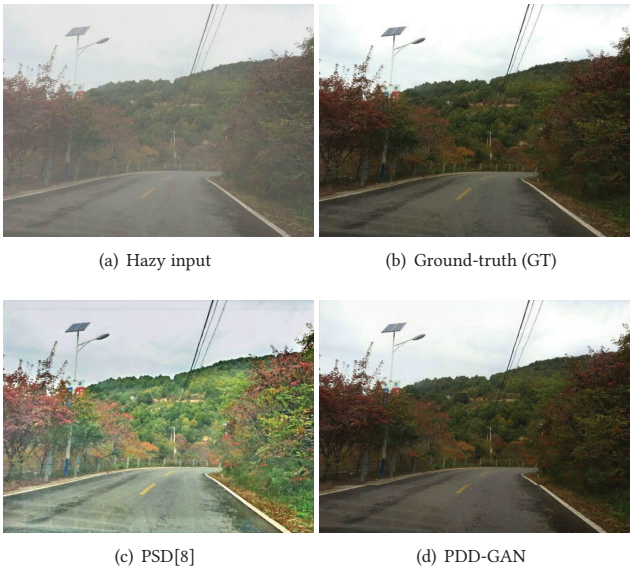
MM '22, October 10–14, 2022, Lisboa, Portugal

ACM ISBN 978-1-4503-9203-7/22/10…$15.00
https://doi.org/10.1145/3503161.3548385

KEYWORDS

single image dehazing, prior-based decoupling, visual representation, contrastive regularization

ACM Reference Format:
Xiaoxuan Chai, Junchi Zhou, Hang Zhou, and Juihsin Lai. 2022. PDD-GAN: Prior-based GAN Network with Decoupling Ability for Single Image Dehazing. In *Proceedings of the 30th ACM International Conference on Multimedia (MM '22), October 10–14, 2022, Lisboa, Portugal.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3503161.3548385



(a) Hazy input (b) Ground-truth (GT)

(c) PSD[8] (d) PDD-GAN

Figure 1: Dehazing examples of two prior-based networks. As shown in the figure, the dehazed image of our model performs better and is closer to the ground-truth (GT).

1 INTRODUCTION

Under poor weather circumstances such as fog, haze and smog, there are extensive tiny suspended particles in the outdoor air, which will refract or scatter the light[10, 45]. These lights mixed with the observed targets lead to degradation in both the structure

and vision of the images captured by outdoor equipment. Consequently, images suffer from definition degradation, color distortion, edge occlusion and texture blur[1, 9, 20]. To address these issues, a number of dehazing algorithms have been developed, and the results can optimize a variety of downstream computer vision tasks, such as semantic segmentation, detection and super resolution.

An important physical scattering model[29, 30] is widely used to describe the formation of hazy images:

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

where, $I(x)$ is an observed hazy image and $J(x)$ is a haze-free image; $t(x)$, x and A represent the medium transmission map, pixel coordinate and global atmospheric light, respectively. Early methods[2, 3, 21, 25, 33, 39, 50] attempted to estimate the transmission map through prior knowledge, and then restore the image through the physical model. However, such physical methods still have inevitable limitations, which strongly depend on estimating the parameters of the physical model and are inapplicable to complex and variable natural scenes. For instance, image dehazing based on dark channel prior[16] is generally not effective for water surface areas and large sky areas. Recovering to a haze-free image only based on this physical model is weakly adaptable because of the difficulty in estimating all parameters properly.

In recent years, deep learning has achieved great success and convolutional neural networks (CNNs) have been introduced into image dehazing. CNN-based methods inversely calculate the parameters of the physical model for dehazing[4, 28, 34, 47]. Usually, it is difficult to accurately estimate these intermediate parameters due to the lack of real-world hazy images and effective physical priors. With the development of image dehazing, a variety of end-to-end CNN-based methods[6, 11, 22, 26, 32, 44, 49] have been proposed to simplify the dehazing problem by learning the hazy-to-clear image conversion directly from the network. Nevertheless, the following problems still exist: (1) Despite being better than traditional methods, the results have defects in edge and color restoration. (2) In addition, most current CNN-based methods just increase the depth of the network. In other words, they enhance the nonlinearity of the system without taking into account the coupling effect of the haze and image background, which lacks the effective compensation of the dehazing process.

To deal with the problem of unclear objects recovery caused by current image dehazing algorithms, Dong *et al.*[13] and Shyam *et al.*[36] proposed to separate the low-frequency and high-frequency information from the image and integrate the frequency prior into the discriminator to fuse the features, failing to improve the ability of the generator to extract features and explain the relationship between the hazy medium and image background. As for color distortion, the dark channel prior[16] and the color attenuation prior[50] conclude that haze will affect the color, saturation and brightness of the image, and a strong coupling exists between them. On the one hand, edge and texture information are blurred by haze, resulting in unclear outlines. On the other hand, the haze obscures some color information and weakens the saturation and brightness of the image. Based on the above considerations, we propose a method to decouple the haze and background on the basis of prior knowledge, in terms of both image structure and vision.

Structural prior knowledge is exploited by leveraging the four-channel image ($\mathrm{RGBF_{HL}}$) as the input and label in our model. Specifically, the band-stop filtering result of the image is used as the fourth channel on the basis of the original three-channel data (RGB). In addition, a contrastive prior loss in hue, saturation and value (HSV) space is designed to achieve visual decoupling in our model, in order to distinguish the difference in color, saturation, and brightness between hazy and clear images. In this paper, we propose a novel Prior-based GAN Network with Decoupling Ability for Single Image Dehazing (denoted as PDD-GAN).

In summary, our contributions are as follows:

- To the best of our knowledge, our model is the first to decouple and compensate hazy images based on structural and visual priors. The dehazing framework can restore hazy images to high-quality haze-free ones, showing state-of-the-art performance in accessible publicly dehazing scenes.
- In addition to the RGB channels, the image information from the band-stop filtering process is included in our input and added as the fourth channel into the network for training. Besides, we have verified that the high and low frequency can preserve more edge and internal information.
- We utilize the distortion of hazy images in color, saturation, and brightness to design a prior loss in HSV space. This loss is regularized by contrastive learning to achieve the visual decoupling of the image.
- We implement sufficient experiments and demonstrate on several image quality evaluation metrics that the four-channel input method and the prior loss method can improve the dehazing effect.

2 RELATED WORK

In recent decades, a lot of dehazing methods have been proposed, which can be roughly divided into three types: physical methods, neural network-based methods and prior-based CNN methods.

2.1 Physical Methods

Early physical methods estimate the parameters of the physical scattering model based on the statistical observations of images. These methods statistically summarized the differences in color (dark channel prior[16] color-lines prior[33], color attenuation prior[50], haze lines prior[3]) and structure (change of detail prior[25], gradient channel prior[21]). For example, Fattal *et al.*[33] observed a generic regularity in natural images where the pixels of small image patches typically exhibited an ID distribution in RGB color space and then used it for recovering the scene transmission. Kaur *et al.*[21] utilized image gradients to estimate depth information and atmospheric light, overcoming problems like texture distortion and transmission map estimation errors to a certain extent. Despite having achieved some success, approaches based on prior knowledge have significant limitations and are usually inapplicable to complex and changeable natural scenes.

2.2 Neural Network-based Methods

Gradually, methods based on neural network were introduced into the dehazing task. Initially, Cai *et al.*[5] and He *et al.*[17] applied the CNN to the dehazing scene by estimating the transmittance

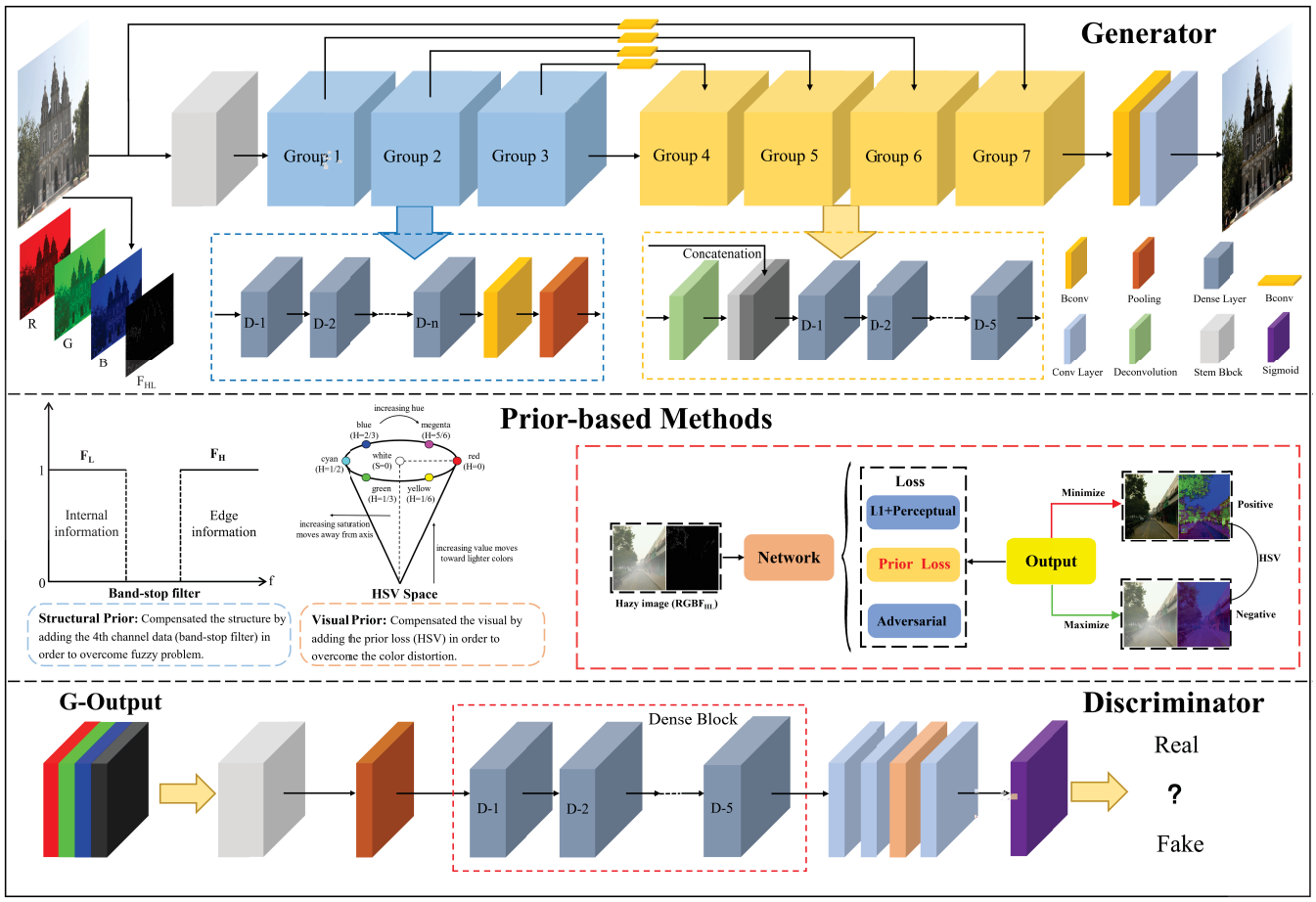
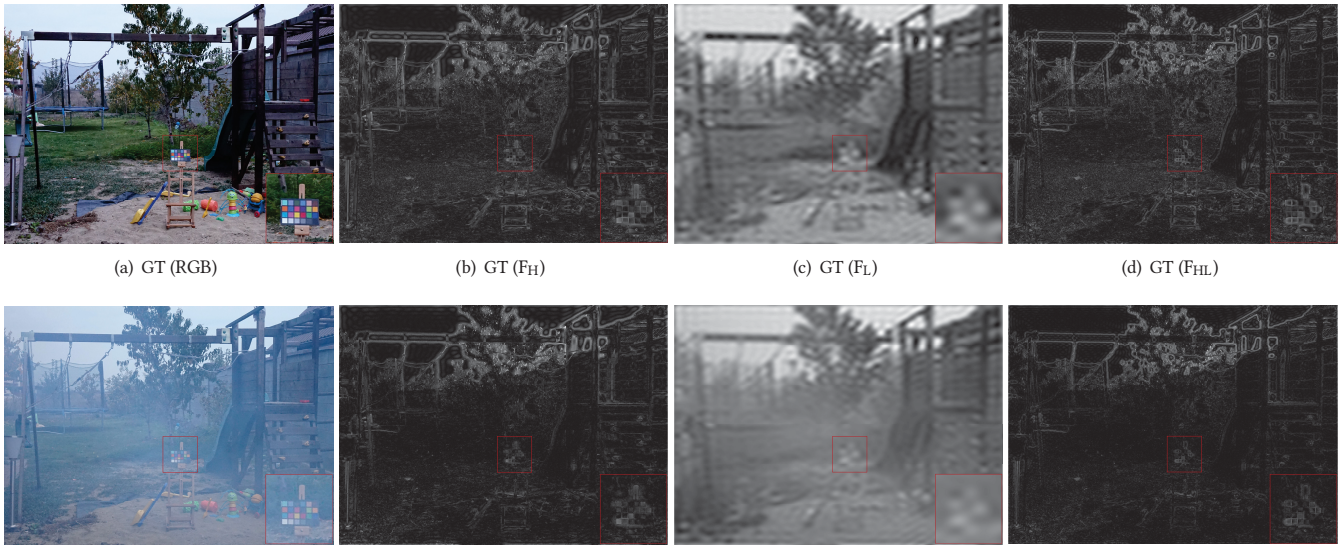


Figure 2: Overview of the proposed PDD-GAN framework. Our model consists of a generator and a discriminator. Prior-based methods include four-channel inputs (RGB channels and F_{HL} channel processed by the band-stop filter) and contrastive prior loss in HSV space.

parameters of the physical model. Ren *et al.*[34] proposed a multi-scale deep neural network to dehaze a single image by learning the relationship between hazy images and their transmission maps. With the development of deep learning in the field of image dehazing, Li *et al.*[22] directly generated clear images through a light-weight end-to-end network AOD-Net, rather than estimate the transmission matrix and atmospheric light separately, as in most previous models. Liu *et al.*[26] implemented an attention-based multi-scale estimation end-to-end network GridDehazeNet, which could dehaze by extracting information between different scales. Xu *et al.*[32] proposed an end-to-end feature fusion attention network FFA-Net, which improved the representation ability through a designed attention module. Dong *et al.*[12] developed a simple and effective enhancement decoder based on the principle of gain and error feedback to restore images. However, dehazing will have an upper limit since the neural network is a black box. These methods fail to take effective priors into consideration, leading to blurred reconstructed images, limited dehazing ability and easily producing color distortion.

2.3 Prior-based CNN Methods

In general, although the physical dehazing algorithms are convenient and this process increases little noise, the scene is severely limited. To achieve a better dehazing effect, some researchers recently combined prior knowledge and CNNs together as a new type dehazing method. Liu *et al.*[13] designed an end-to-end GAN network, adding frequency information as additional prior knowledge to the discriminator for fusing features. Shyam *et al.*[36] proposed a dual discriminator that independently learned low-frequency and high-frequency information to guarantee the restoration of image color and structural properties. Chen *et al.*[8] designed a synthetic-to-realistic dehazing framework via a loss committee with physical priors. However, these methods only take some priors as additional constraints and can't utilize them effectively. Therefore, a prior-based end-to-end dehazing GAN network have been proposed in this paper.



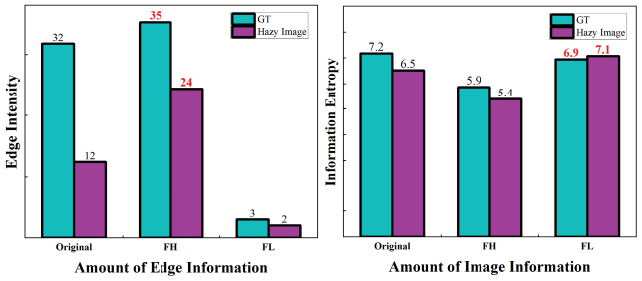
(a) GT (RGB) (b) GT (F_H) (c) GT (F_L) (d) GT (F_{HL})

(e) Hazy Image (RGB) (f) Hazy Image (F_H) (g) Hazy Image (F_L) (h) Hazy Image (F_{HL})

Figure 3: High frequency (F_H), low frequency (F_L) and band-stop filtering (F_{HL}) results of GT and hazy image. As shown in the figure, F_H mainly contains outlines of the object. F_L retains internal information without the texture. And F_{HL} is similar to F_H but retains more details.

3 METHODOLOGY

To equip the network with decoupling ability and avoid learning the same features multiple times, we introduce two kinds of priors into our framework. The first one transfers images to the frequency domain for structural decoupling and the second one analyzes the loss in HSV space for visual recovery.



(a) Edge-intensity (b) Information-entropy

Figure 4: Statistical chart of high frequency and low frequency information. (a) is the edge intensity calculated by sobel operator of hazy images and GT[35, 48]. (b) is the information entropy of images calculated by entropy formula[31].

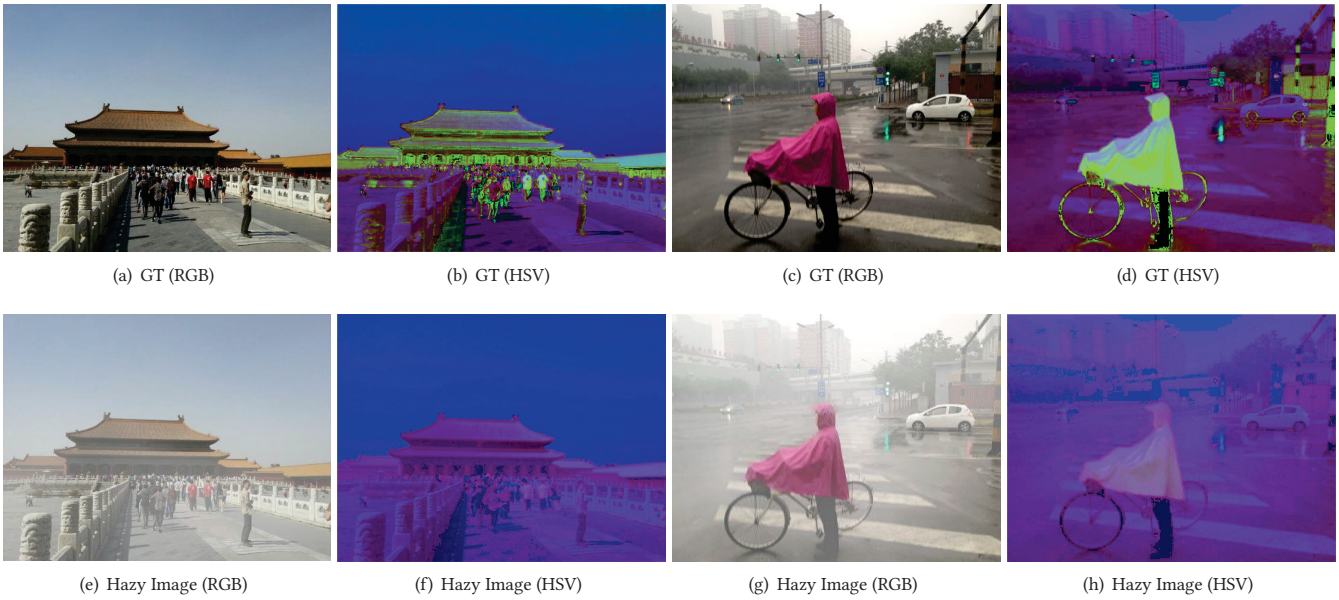
3.1 Prior Knowledge Decoupling

It can be known from some priors [3, 16, 21, 25, 33, 50] that the haze or fog is blended with the image background in various ways. Particularly, the object's outlines are blurred, and the hue, saturation and lightness are significantly affected by the haze. Compared with the ground-truth (GT) in structure, the texture or edge of the hazy image can hardly be distinguished. Additionally, the haze visually drops some color details and weakens the saturation and brightness of the image. If these issues are solved unilaterally only by increasing the depth of the neural network, models will be easily over fitting. Thus, the following is analyzed in terms of both structural and visual decoupling to overcome the haze-split challenge.

Structural Decoupling. The structural information of an image can be refined into a collection of edge and internal information. The former refers to the set of pixel points with step change in the gray scale of the surrounding pixels, namely the outline of the image, while the latter is the collection of those pixel points where image pixels change slowly in gray scale, and denotes the main content contained in the image.

In computer vision and signal processing, the structural information of an image can be transformed from the spatial domain to the frequency one. Frequency is a basic indicator characterizing the gradient of gray scale, specifically the sharpness of gray scale changes in an image. High-frequency information (F_H) refers to dramatically changing parts, like outlines and delicate textures. We calculate the edge intensity of hazy and haze-free images processed by high-pass filter and low-pass filter respectively on the dataset NTIRE'18[40]. As shown in Fig. 4(a), the values of F_H are the highest (GT: 35, hazy image: 24) and the gap between GT and hazy images is smaller than the origin (20→11). High value means the ability of F_H to retain more edge information, and small gap means less haze influence, demonstrating that F_H images eliminating haze disturbance indeed preserve edge information.

On the contrary, low-frequency information (F_L) refers to slowly changing parts, laying more emphasis on the primary information



(a) GT (RGB) (b) GT (HSV) (c) GT (RGB) (d) GT (HSV)

(e) Hazy Image (RGB) (f) Hazy Image (HSV) (g) Hazy Image (RGB) (h) Hazy Image (HSV)

Figure 5: Two examples of hazy and haze-free images transformed into HSV images. Compared to RGB, the HSV images shows a great visual difference between the hazy images and GT.

of the image, such as brightness, color and saturation information. We calculate the information entropy on the same data set to measure the information retained in the image. As shown in Fig. 4(b), the values of F_L are 6.9 and 7.1 of GT and hazy images respectively. The gap between GT and hazy images is smaller than the origin (0.7→0.2). High value means the ability of F_L retaining more internal information, and small gap means less haze influence, demonstrating that F_L images eliminating the haze disturbance indeed preserve internal information.

Based on what is discussed above, we use the advantages of high and low frequency by band-stop filtering. Some filtering examples in Fig. 3. The objects' outlines of the F_H images are clear but drop the internal information. The main parts are retained in F_L images while outlines and texture are blurred. Our network takes frequency prior as additional constraints to compensate edge information and achieve the structural decoupling of the model. Specifically, after the Fourier transform of the image, we process the outcome with a band-stop filter to generate the fourth channel image ($RGBF_{HL}$).

Visual Decoupling. Images can be visually represented in multiple spaces such as RGB, HSV and HSL. Different color spaces emphasize different information. For example, RGB space[14] is more suitable for image processing at the hardware level, and the relationship is not intuitive between the values of the three color components and generated colors, making quantitative analysis difficult. However, HSV space[38] is more similar to the human perception of color information which contains the brightness and saturation of an image. Therefore, transforming images into a more appropriate space can better represent visual information.

The difference between hazy and haze-free images can be clearly distinguished in a visual way. Firstly, He *et al.*[16] found a key observation that a majority of local patches in haze-free outdoor images contain some pixels with very low intensities in at least one color channel. Secondly, Zhu *et al.*[50] claimed that the concentration of haze is proportional to the difference between brightness and saturation. Beyond that, most prior knowledge states that haze has an influence on the color, saturation and brightness of an image, and they are coupled tightly together. We convert the image to HSV space due to the difficulty in separating these differences in RGB space. Fig. 5 shows a huge visual difference between the hazy and haze-free images. Thus a prior loss is designed to achieve visual representation. This section is highlighted in 3.3 Loss Function.

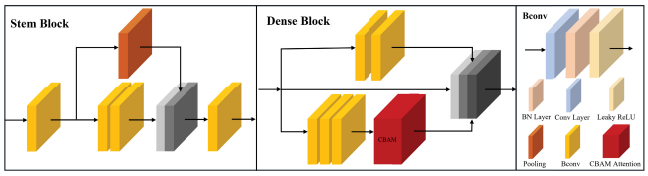
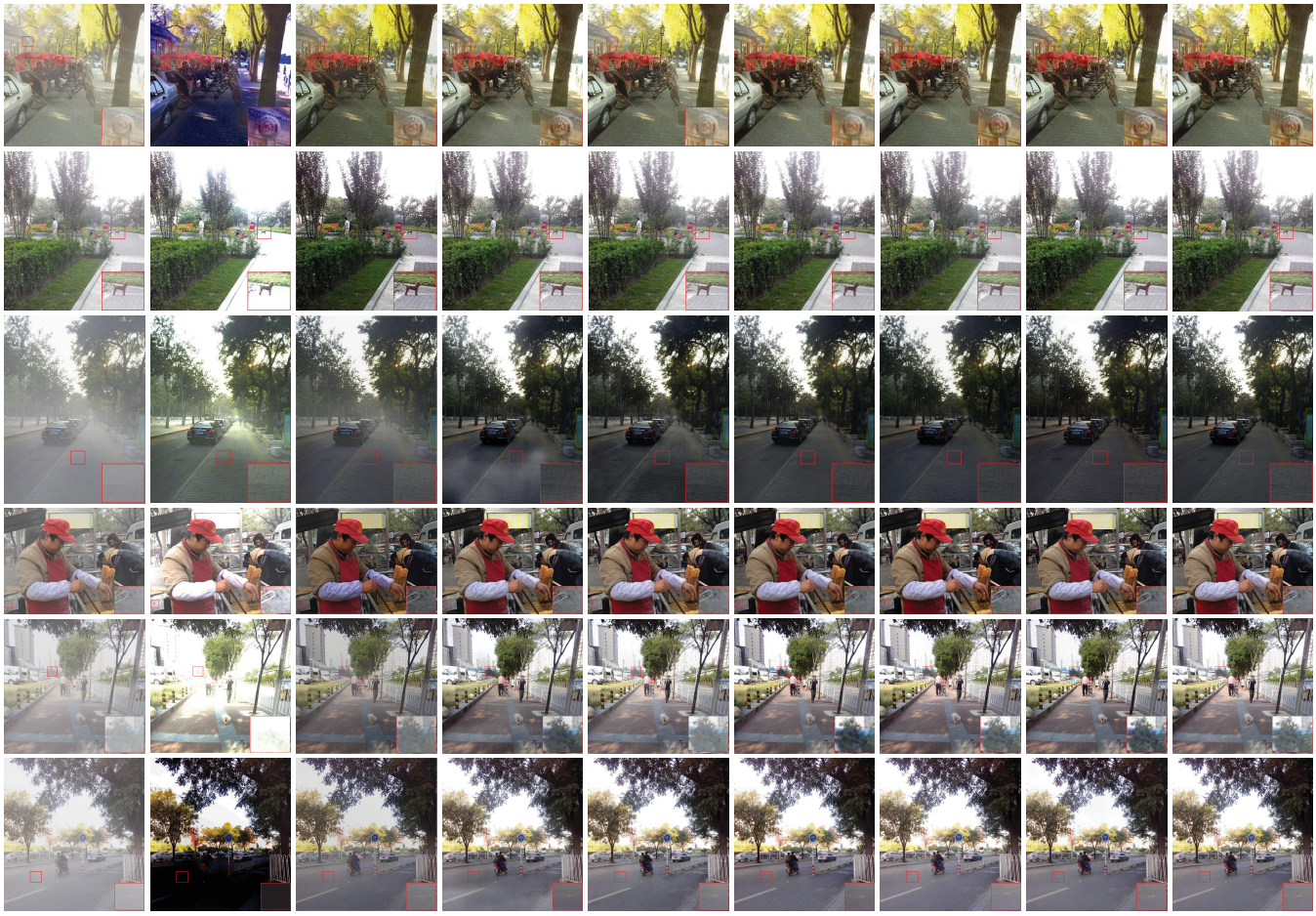


Figure 6: The structure of Stem Block and Dense Block

3.2 Network Architecture

The architecture of PDD-GAN is introduced in this section. As shown in Fig. 2, the lightweight PeleeNet[41] is used as the backbone of the network. First, input is processed by stem block for shallow feature extraction. Then, feature maps pass through three down-sample groups (1-3) which composed of 4, 6, 8 dense blocks respectively to extract deep features. After that, the generated feature maps are upsampled by groups (4-7) containing five dense blocks. During the upsampling process, output features are fused



(a) Inputs (b) Dehaze-Net (c) AOD-Net (d) GridDehazeNet (e) FFA-Net (f) MSBDN-DFF (g) DehazeFlow (h) PDD-GAN (i) GT

Figure 7: Comparison of PDD-GAN with the state-of-the-art methods on SOTS outdoor benchmark.

with the downsampled ones after four jump connections and finally recovered to a haze-free (RGBF_{HL}) image.

What's more, the CBAM attention module[43] in the dense block is deployed to make the model pay more attention on channels, as shown in Fig. 6. In this way, in addition to ensure the feature representation ability, the network would increase the weight of the fourth channel decoupling information.

3.3 Loss Function

L1 loss. To measure the pixel-wise reconstruction effect, we use the L1 loss to calculate the loss on four channels (RGBF_{HL}).

$$\mathcal{L}^{L1} = \frac{1}{N}\sum_{i=1}^{N}\left\|G_{4c}(I_{4c}^{i}) - J_{4c}^{i}\right\| \tag{2}$$

Where I_{4c}^{i} and J_{4c}^{i} represent the hazy image and corresponding GT with four channels, respectively. $G_{4c}(I_{4c}^{i})$ is the four-channel output produced by generator.

Perceptual loss. We adopt perceptual loss[19] to evaluate the semantic-wise reconstruction effect, including perceptual and style similarity in feature map space.

$$\mathcal{L}^{Pl} = \frac{1}{N}\sum_{i=1}^{N}\left(\left\|\phi(I_{3c}^{i}) - \phi(J_{3c}^{i})\right\|_2^2 + \left\|\varphi_1(I_{3c}^{i}) - \varphi_1(J_{3c}^{i})\right\|_2^2\right) \tag{3}$$

$$\varphi_1(x) = \phi(x)\cdot\phi(x') \tag{4}$$

Where I_{3c}^{i} and J_{3c}^{i} denote the hazy and clear images with RGB channels, respectively. $\phi(\cdot)$ represents the 30th layer feature map of vgg16 network[37], x' is the transpose of x .

Prior loss. According to the difference between hazy and clear images in HSV space, we design a prior-based loss function. This prior loss can compare the difference between the output and GT in hue, saturation and brightness.

$$\mathcal{L}^{Pr} = \frac{1}{N}\sum_{i=1}^{N}\left(\left\|G_h(I_{4c}^{i}) - J_h^{i}\right\| + \left\|\frac{\varphi_2(G_{sv}(I_{4c}^{i})) - \varphi_2(J_{sv}^{i})}{\varphi_2(G_{sv}(I_{4c}^{i})) - \varphi_2(I_{sv}^{i}) + \varepsilon}\right\|\right) \tag{5}$$

$$\varphi_2(x) = x_s - x_v \tag{6}$$

where $G_h(I^i_{4c})$ and $G_{sv}(I^i_{4c})$ are the outputs with H channel and S, V channels of the generator, respectively. J^i_h and J^i_{sv} denote H channel and S, V channels of GT, respectively. I^i_{sv} is the S and V channels of the hazy image. ε is a minimal constant to avoid the meaninglessness of the denominator. x_s and x_v represent the S and V channel of input x, respectively.

Based on the linear relationship between haze and the value difference between saturation and brightness[50], we propose a contrast regularization method[7, 15], and regard the hazy images as negative samples and GT as positive samples. The purpose is to expand the distance between dehazed and hazy images and reduce the gap between dehazed images and GT as well. Contrast regularization draws on the idea of clustering method, which makes the network learn the coupling characteristics between haze and image background by itself. The negative samples, hazy images, are used to restrict the feasible solution space, and the ablation study demonstrates prior loss improves the dehazing performance (see 4.3).

Adversarial loss. $D(G(I^i_{4c}))$ is the output of the discriminator.

$$\mathcal{L}^{Al} = \frac{1}{M}\sum_{i=1}^{M} \log\left(1 - D(G(I^i_{4c}))\right) \tag{7}$$

Loss of the generator. We combine the loss mentioned above together as generator loss function to regularize our GAN network.

$$\mathcal{L}^{Gr} = \lambda_1\mathcal{L}^{L1} + \lambda_2\mathcal{L}^{Pl} + \lambda_3\mathcal{L}^{\mathrm{Pr}} + \lambda_4\mathcal{L}^{Al} \tag{8}$$

where λ_1, λ_2, λ_3 and λ_4 are weights of those sub-loss functions.

4 EXPERIMENTS

In this section, we make a comparison with seven state-of-the-art models, and then ablation study is complemented to verify the effectiveness of the proposed prior-based decoupling method.

4.1 Experiment Settings

Datasets. Our model is trained and verified on the RESIDE dataset[23]. The dataset contains RESIDE-V0, RESIDE-Standard and RESIDE-β three versions. All outdoor training sets (OTS) are chosen for training, including OTS in RESIDE-V0 and data in RESIDE-β. The first set contains 313,950 outdoor hazy and clean image pairs, and the second set contains 72,135 image pairs. The number of image pairs is 386,085 in total. We evaluate our model and compare the results with the state-of-the-art methods on the OTS of the SOTS benchmark provided by RESIDE.

Training Details. We use the ideal band-stop filter which set band width as 60HZ and the banned distance from the center point is 25HZ in Fourier space. In order to augment the data, the images are randomly cropped into 224×224 patches as training data, with a random horizontal flip. The model is optimized by Adamw optimizer[27] with an initial learning rate of 0.0001, β_1=0.9, β_2=0.999 and decay rate of 1×10^{-4} for both generator and discriminator. The parameters of the generator are set as (λ_1, λ_2, λ_3, λ_4) = (1.0, 0.5, 1.0, 0.5).

Evaluation Metrics. To quantify and compare the performance of the model, we select two objective evaluation metrics: Peak Signal to Noise Ratio (PSNR)[18], and Structural Similarity index (SSIM)[42]. Given a dehazed result and corresponding GT, PSNR measures their average pixel similarity, and SSIM measures their structural similarity. The model PSD[8] provided no PSNR and SSIM metrics on the testing dataset. Hence, we use their trained checkpoint and then test metrics by ourselves.

4.2 Comparisons with State-of-the-art Methods

To illustrate the effectiveness and generalization ability of our model, we make a comparison with seven state-of-the-art models, including DehazeNet[5], AOD-Net[22], GridDehazeNet[26], FFA-Net[32], MSBDN-DFF[12], PSD[8] and DehazeFlow[24]. As shown in Tab. 1, the performance of PDD-GAN and SOTA methods is summarized on the SOTS outdoor test set, a subset of the RESIDE dataset. Compared with others, our PDD-GAN achieves the best performance with 35.1dB PSNR and 0.9897 SSIM.

Also, PDD-GAN is compared with SOTA methods on the quality of restored images, as shown in Fig. 7. It can be observed that DehazeNet and AOD-Net cannot successfully remove dense haze, and the color is disturbed significantly by the haze (see Fig. 7(b), (c)). Although DehazeFlow and FFA-Net can restore high quality dehazing results, low brightness and fuzzy problem appear as well (see Fig. 7(e), (g)). As shown in Fig. 1, we compare our model with another prior-based method PSD. PSD just uses the physical prior to fine-tune the dehazing process, which fails to effectively decouple the haze and suffers from the color distortion. More importantly, our results recover better visually, which can restore hazy images into more natural ones with similar patterns to GT both in structure and vision.

Table 1: Quantitative comparisons with other state-of-the-art methods on the OTS of SOTS datasets.

Method	PSNR(SOTS)	SSIM(SOTS)
(TIP'16) Dehaze-Net	24.75	0.9269
(ICCV'17) AOD-Net	24.14	0.9198
(ICCV'19)GridDehazeNet	30.86	0.9819
(AAAI'20) FFA-Net	33.57	0.9840
(CVPR'20) MSBDN-DFF	33.79	0.9840
(CVPR'21) PSD	21.85	0.9436
(ACMM'21) DehazeFlow	34.24	0.9849
PDD-GAN(Ours)	**35.10**	**0.9897**

Table 2: Comparisons on SOTS for different channel configurations. The number of ★ represents the importance of metrics.

Input	PSNR★★★ (SOTS/NTIRE'18)	SSIM★★ (SOTS/NTIRE'18)	FSIM★ (SOTS/NTIRE'18)
RGB	26.3452/17.4129	0.9587/0.6541	0.9882/0.8047
$RGBF_L$	26.3135/18.0533	0.9579/0.6762	0.9876/0.8073
$RGBF_H$	27.0237/18.5934	0.9618/0.6559	0.9888/0.7873
$RGBF_{HL}$	**27.3425/18.8735**	**0.9652/0.6781**	**0.98887/0.8112**

4.3 Ablation Study

Two ablation experiments are designed to better demonstrate that our structure and vision prior-based methods have decoupling ability for image dehazing. First, we use RESIDE dataset (ITS and OTS) with a total of 86,125 images to train our model for 20 epochs, and we evaluate our model on SOTS and NTIRE'18 tset sets. The images are processed into four types: RGB, RGBF_L, RGBF_H and RGBF_HL. As shown in Tab. 2, the PSNR, SSIM and FSIM (feature similarity)[46] image metrics of the input RGBF_HL get the highest score on both testing sets. Compared to RGB, RGBF_HL achieves higher PSNR, SSIM and FSIM with the gains of 0.997/1.461dB, 0.0065/0.024 and 0.0006/0.0065 respectively and RGBF_H improves the metrics as well. Although RGBF_L gets lower score on SOTS, the combination of F_L and F_H performs better.

The second experiment follows two control groups: (1) with prior loss ($\mathcal{L}^{Gr}$), (2) without prior loss ($\mathcal{L}'$ as in Formula. 9). Our model is trained on training set (OTS) with RGBF_HL inputs and evaluated on SOTS, as shown in Tab. 3. It can be seen that adding a prior loss to our model can improve the final dehazing results and reconstruction performance, and the PSNR increases by 2.0db, SSIM increases by 0.031 and FSIM escalates by 0.009.

$$\mathcal{L}'=\lambda_1\mathcal{L}^{L1}+\lambda_2\mathcal{L}^{Pl}+\lambda_4\mathcal{L}^{Al} \tag{9}$$

Finally, the experiments show that the RGBF_HL input and prior loss can indeed improve the dehazing performance and achieve the decoupling ability.

Table 3: Ablation study results. By using prior loss, our method obtains significant improvements.

Loss Function	PSNR↑	SSIM↑	FSIM↑
$\mathcal{L}'$	33.10	0.9587	0.9882
$\mathcal{L}^{Gr}$	**35.10**	**0.9897**	**0.9972**

5 CONCLUSION

In this paper, we have proposed a prior-based GAN network based on PeleeNet with decoupling ability for single image dehazing. Besides, prior-based methods are put forward, including adding the image processed by a band-stop filter as the fourth channel into the network and designing a novel prior loss function in order to offer compensation information to our network. Compared with other state-of-the-art methods, our method achieves the best performance, which can generate high quality images. In addition, other extensive experiments have demonstrated that F_H and F_L images represent edge and internal information, respectively. Owing to the frequency domain and HSV information, the network can better extract the structural and visual features of the image.

6 ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China under Grant 61903215, in part by the Natural Science Foundation of Guangdong Province under Grant 2022A1515010543.

REFERENCES

[1] Z. Anvari and V. Athitsos. 2020. Dehaze-GLCGAN: Unpaired Single Image Dehazing via Adversarial Training. (2020).
[2] D. Berman, T. Treibitz, and S. Avidan. 2016. Non-local Image Dehazing. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
[3] Dana Berman, Tali Treibitz, and Shai Avidan. 2020. Single Image Dehazing Using Haze-Lines. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42, 3 (2020), 720–734. https://doi.org/10.1109/TPAMI.2018.2882478
[4] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao. 2016. DehazeNet: An End-to-End System for Single Image Haze Removal. *IEEE Transactions on Image Processing* 25, 11 (2016), 5187–5198.
[5] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao. 2016. DehazeNet: An End-to-End System for Single Image Haze Removal. *IEEE Transactions on Image Processing* 25, 11 (2016), 5187–5198.
[6] C. M. Chang, C. S. Sung, and T. N. Lin. 2021. DAMix: Density-Aware Data Augmentation for Unsupervised Domain Adaptation on Single Image Dehazing. (2021).
[7] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton. 2020. A Simple Framework for Contrastive Learning of Visual Representations. (2020).
[8] Zeyuan Chen, Yangchao Wang, Yang Yang, and Dong Liu. 2021. PSD: Principled Synthetic-to-Real Dehazing Guided by Physical Priors. 7176–7185. https://doi.org/10.1109/CVPR46437.2021.00710
[9] Xiaofeng Cong, Jie Gui, Kaichao Miao, Jun Zhang, Bing Wang, and Peng Chen. 2020. Discrete Haze Level Dehazing Network. In *MM '20: The 28th ACM International Conference on Multimedia*.
[10] S. D. Das and S. Dutta. 2020. Fast Deep Multi-patch Hierarchical Network for Nonhomogeneous Image Dehazing. (2020).
[11] Q. Deng, Z. Huang, C. C. Tsai, and C. W. Lin. 2020. *HardGAN: A Haze-Aware Representation Distillation GAN for Single Image Dehazing*. Computer Vision – ECCV 2020.
[12] H. Dong, J. Pan, L. Xiang, Z. Hu, and M. H. Yang. 2020. Multi-Scale Boosted Dehazing Network with Dense Feature Fusion. *arXiv* (2020).
[13] Y Dong, Y Liu, H Zhang, S. Chen, and Y Qiao. 2020. FD-GAN: Generative Adversarial Networks with Fusion-Discriminator for Single Image Dehazing. *Proceedings of the AAAI Conference on Artificial Intelligence* 34, 7 (2020), 10729–10736.
[14] RC Gonzales and RE Woods. 2018. Digital image processing 4th edition.
[15] K. He, H. Fan, Y. Wu, S. Xie, and R Girshick. 2019. Momentum Contrast for Unsupervised Visual Representation Learning. (2019).
[16] Kaiming He, Jian Sun, and Xiaoou Tang. 2011. Single Image Haze Removal Using Dark Channel Prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, 12 (2011), 2341–2353. https://doi.org/10.1109/TPAMI.2010.168
[17] Z. He, V. Sindagi, and V. M. Patel. 2017. Joint Transmission Map Estimation and Dehazing using Deep Networks. *IEEE Transactions on Circuits and Systems for Video Technology* PP, 99 (2017).
[18] Q. Huynh-Thu and M. Ghanbari. 2008. Scope of validity of PSNR in image/video quality assessment. *Electronics Letters* 44, 13 (2008), 800–801.
[19] J. Johnson, A. Alahi, and L. Fei-Fei. 2016. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. *Springer, Cham* (2016).
[20] A. Kar, S. K. Dhara, D. Sen, and P. K. Biswas. 2020. Transmission Map and Atmospheric Light Guided Iterative Updater Network for Single Image Dehazing. (2020).
[21] M. Kaur, D. Singh, V. Chahar, and K. Sun. 2020. Color Image Dehazing using Gradient Channel Prior and Guided L0 Filter. *Information Sciences* 521 (2020).
[22] B. Li, X. Peng, Z. Wang, J. Xu, and F. Dan. 2017. AOD-Net: All-in-One Dehazing Network. In *2017 IEEE International Conference on Computer Vision (ICCV)*.
[23] B. Li, W. Ren, D. Fu, D. Tao, and Z. Wang. 2017. RESIDE: A Benchmark for Single Image Dehazing. (2017).
[24] Hongyu Li, Jia Li, Dong Zhao, and Long Xu. 2021. DehazeFlow: Multi-scale Conditional Flow Network for Single Image Dehazing. (2021).
[25] J. Li, H. Zhang, D. Yuan, and M. Sun. 2015. Single image dehazing using the change of detail prior. *Neurocomputing* 156 (2015), 1–11.
[26] X Liu, Y. Ma, Z. Shi, and J. Chen. 2020. GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*.
[27] I. Loshchilov and F. Hutter. 2017. Decoupled Weight Decay Regularization. (2017).
[28] Huimin Lu, Yujie Li, Shota Nakashima, and Seiichi Serikawa. 2015. Single image dehazing through improved atmospheric light estimation. *Multimedia Tools and Applications* 75, 24 (oct 2015), 17081–17096. https://doi.org/10.1007/s11042-015-2977-7
[29] McCartney, E., J., Hall, Freeman, and F. 1977. Optics of the Atmosphere: Scattering by Molecules and Particles. *Phys Today* (1977).
[30] S. G. Narasimhan and S. K. Nayar. 2002. Vision and the Atmosphere. *International Journal of Computer Vision* 48, 3 (2002), 233–254.
[31] J. A. Núez, P. M. Cincotta, and F. C. Wachlin. 1996. Information entropy. *Celestial Mechanics Dynamical Astronomy* 64, 1-2 (1996), 43–53.

[32] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Jia. 2019. FFA-Net: Feature Fusion Attention Network for Single Image Dehazing. (2019).
[33] Raanan and Fattal. 2014. Dehazing Using Color-Lines. *ACM Transactions on Graphics (TOG)* 34, 1 (2014).
[34] W. Ren, L. Si, Z. Hua, J. Pan, and M. H. Yang. 2016. Single Image Dehazing via Multi-scale Convolutional Neural Networks. *Springer, Cham* (2016).
[35] John C. Russ. 2011. *The Image Processing Handbook, Sixth Edition* (6th ed.). CRC Press, Inc., USA. 192 pages.
[36] P. Shyam, K. J. Yoon, and K. S. Kim. 2021. Towards Domain Invariant Single Image Dehazing. (2021).
[37] K. Simonyan and A. Zisserman. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. *Computer Science* (2014).
[38] Slawomir Skoneczny. 2012. Nonlinear image sharpening in the HSV color space. *Przeglad Elektrotechniczny* 88, 2 (2012), 140–144.
[39] R. T. Tan. 2008. Visibility in bad weather from a single image. In *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), 24-26 June 2008, Anchorage, Alaska, USA*.
[40] P. F. Wang. 2018. NTIRE 2018 Challenge on Image Dehazing: Methods and Results. (2018).
[41] Robert J Wang, Xiang Li, and Charles X Ling. 2018. Pelee: A Real-Time Object Detection System on Mobile Devices. In *Advances in Neural Information Processing Systems 31*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Eds.). Curran Associates, Inc., 1967–1976. http://papers.nips.cc/paper/7466-pelee-a-real-time-object-detection-system-on-mobile-devices.pdf
[42] Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, and Eero P. Simoncelli. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process* 13, 4 (2004).
[43] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon. 2018. CBAM: Convolutional Block Attention Module. *Springer, Cham* (2018).
[44] H. Wu, Y Qu, S. Lin, J. Zhou, R. Qiao, Z. Zhang, Y Xie, and L. Ma. 2021. Contrastive Learning for Compact Single Image Dehazing. (2021).
[45] Q. Wu, J. Zhang, W. Ren, W. Zuo, and X. Cao. 2020. Accurate Transmission Estimation for Removing Haze and Noise From a Single Image. *IEEE Transactions on Image Processing* 29 (2020), 2583–2597.
[46] S. Xu, X. Liu, and S. Jiang. 2015. A Fast Feature Similarity Index for Image Quality Assessment. (2015).
[47] H. Zhang and V. M. Patel. 2018. Densely Connected Pyramid Dehazing Network. (2018).
[48] X. Zhang, X. Feng, W. Wang, and W. Xue. 2013. Edge Strength Similarity for Image Quality Assessment. *IEEE Signal Processing Letters* 20, 4 (2013), 319–322.
[49] Dong Zhao, Jia Li, Hongyu Li, and Long Xu. 2021. Complementary Feature Enhanced Network with Vision Transformer for Image Dehazing. https://doi.org/10.48550/ARXIV.2109.07100
[50] Qingsong Zhu, Jiaming Mai, and Ling Shao. 2015. A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior. *IEEE Transactions on Image Processing* 24, 11 (2015), 3522–3533. https://doi.org/10.1109/TIP.2015.2446191