DFR-Net: Density Feature Refinement Network for Image Dehazing Utilizing Haze Density Difference

Zhongze Wang, Haitao Zhao, Lujian Yao, Jingchao Peng, Kaijie Zhao.

Abstract—In image dehazing task, haze density is a key feature and affects the performance of dehazing methods. However, some of the existing methods lack a comparative image to measure densities, and others create intermediate results but lack the exploitation of their density differences, which can facilitate perception of density. To address these deficiencies, we propose a density-aware dehazing method named Density Feature Refinement Network (DFR-Net) that extracts haze density features from density differences and leverages density differences to refine density features. In DFR-Net, we first generate a proposal image that has lower overall density than the hazy input, bringing in global density differences. Additionally, the dehazing residual of the proposal image reflects the level of dehazing performance and provides local density differences that indicate localized hard dehazing or high density areas. Subsequently, we introduce a Global Branch (GB) and a Local Branch (LB) to achieve density-awareness. In GB, we use Siamese networks for feature extraction of hazy inputs and proposal images, and we propose a Global Density Feature Refinement (GDFR) module that can refine features by pushing features with different global densities further away. In LB, we explore local density features from the dehazing residuals between hazy inputs and proposal images and introduce an Intermediate Dehazing Residual Feedforward (IDRF) module to update local features and pull them closer to clear image features. Sufficient experiments demonstrate that the proposed method achieves results beyond the state-of-the-art methods on various datasets.

Index Terms—Image Processing, Image Dehazing, Deep Learning, Density-aware.

I. INTRODUCTION

HAZE is a common atmospheric phenomenon caused by the accumulation of aerosol particles. It can cause severe quality degradation of images, which can affect subsequent computer vision tasks. Therefore, developing effective techniques for haze removal is essential to improve the quality of images and ensure accurate results of downstream tasks [1–3].

After decades of study, researchers [4, 5] model this atmospheric phenomenon as:

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

where $\mathbf{I}(x)$ represents the hazy image; $\mathbf{J}(x)$ represents the clear image; $t(x)$ and $\mathbf{A}$ stand for the transmission map (T-map) and the global atmospheric light separately. And this model is commonly called the atmospheric scattering model (ASM).
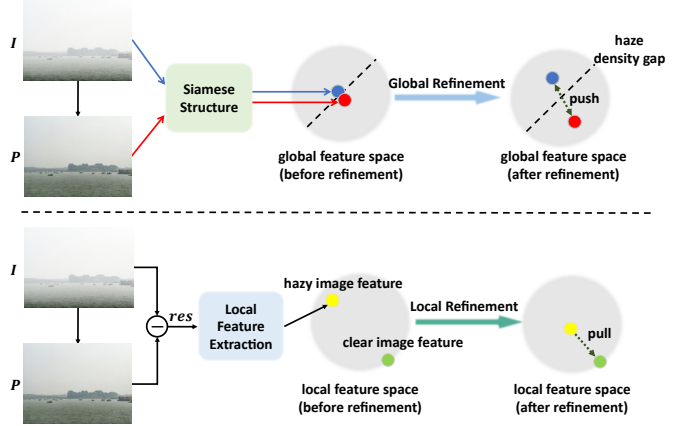
Fig. 1. The main idea of our DFR-Net. By utilizing the density difference between the input (I) and the generated proposal image (P), global and local density features are refined in different ways: features of different global densities are pushed farther apart and local density features are pulled in towards clear image features.

To improve image quality and highlight image details captured in hazy weather, numerous image dehazing methods have been proposed. Prior-based methods [6–10] rely on statistical analysis of haze images and handcrafted priors to recover haze-free images. However, these methods have limitations in their robustness due to their reliance on specific assumptions, which may not hold in different scenes.

With the success of deep neural networks in high-level tasks, data-driven dehazing methods [10–15] have become mainstream. Compared to prior-based methods, deep learning methods demonstrate stronger capabilities in feature extraction and image restoration. However, early deep learning methods neglect the uneven distribution of haze, resulting in redundancy in network design and inefficiency in feature extraction [15]. One idea to improve these methods is to enable the network to learn features about the haze density.

Research works [15–19] have proposed density-aware dehazing methods. Haze density describes the distribution of haze and impacts the effectiveness of a dehazing method. Several methods [16, 18, 19] estimate the T-map to obtain haze density information, which is inversely proportional to the T-map [16]. However, these methods require T-map labels, which can be difficult to obtain. To avoid this problem, methods [15, 17] directly extract density-related features. Nonetheless, these methods still have shortcomings: they lack a comparator to measure density, and the learning process of density features lacks interpretability.

As image dehazing is an ill-posed problem, estimating a

clear image directly from a hazy input using a feedforward network is challenging [20]. Additionally, distinguishing haze densities from a single hazy image is also difficult, and a comparable image that helps the network learn to perceive different haze densities is necessary for density-awareness. While some research works [15, 20, 21] have proposed methods that generate intermediate images to alleviate this problem, they do not take the density differences between the intermediate results and the hazy inputs into much consideration, limiting the network's ability to fully perceive haze density.

To address this issue, we propose a Density Feature Refinement Network (DFR-Net) that achieves density awareness by utilizing density difference. To fully explore haze density feature, we first generate a dehazing proposal image ($\mathbf{P}$, as shown in Fig. 1), that can provide a comparison of density information from both global and local perspectives. Our proposal image is generated by a simple U-net structure and has an overall lower haze density than the haze input ($\mathbf{I}$), which brings global density differences. The different dehazing performance of different areas reflected in the dehazing residuals ($\mathbf{res}$) leads to differences in local density and hints to areas with haze removal challenges. Subsequently, we propose a 2-branch structure consisting Global Branch (GB) and Local Branch (LB) *based on the density difference information between $\mathbf{P}$ and $\mathbf{I}$* to comprehensively extract density features and refine them in different ways.

In detail, the scenes are the same between $\mathbf{P}$ and $\mathbf{I}$, differing only in haze density. Therefore, we propose to use Siamese structures for feature extraction for $\mathbf{P}$ and $\mathbf{I}$, enabling the network to understand both the different densities in one forward process. In addition, considering that the two features contain part of the same information, to push them farther away and highlight the density information, we design a Global Density Feature Refinement (GDFR) module to refine the features. Locally, the dehazing residual between $\mathbf{P}$ and $\mathbf{I}$ contains hints of the local density information. The dehazing residual represents the performance of dehazing and areas with small residual values tend to be more heavy or hard dehazing regions. Hence, we propose to learn local density features from the dehazing residual ($\mathbf{P} - \mathbf{I}$) in a split and merge (S&M) way. To refine local features, several Intermediate Dehazing Residual Feedforward (IDRF) modules are used, which can pull local features closer to clear image features. The refinement illustration is exemplified in Fig.1. Both the two branches give a predicted haze-free image and we perform an adaptive fusion on them to gain the final dehazing result.

Compared to other density-aware methods, on the one hand, our DFR-Net does not rely on additional T-map annotations and does not predict the T-map, thus reducing the manual workload and the possible loss of information due to the T-map prediction process [22]. On the other hand, unlike the way other methods extract density features, e.g., PMNet extracts density features from the splicing of hazy inputs and pseudo-haze-free images via an SHA module, we analyze the relationship of haze density between $\mathbf{P}$ and $\mathbf{I}$, and design a network with interpretability to extract and refine density features from the density difference. This makes our DFR-Net contain density-related prior knowledge.
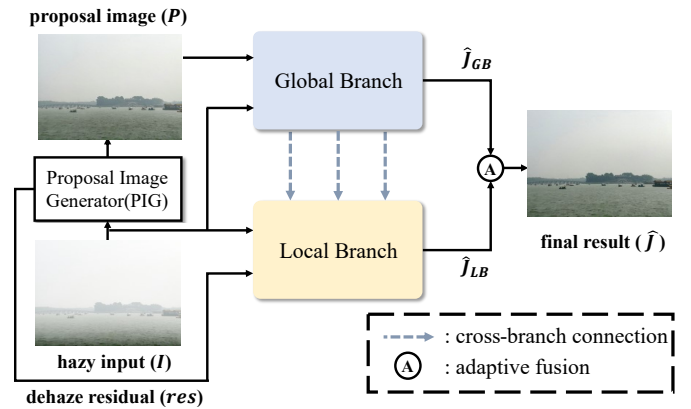

Fig. 2. The pipeline of DFR-Net. DFR-Net first generates a proposal image ($\mathbf{P}$) by PIG, and $\mathbf{P}$ is input to the subsequent two-branch network together with the hazy input ($\mathbf{I}$). Each branch predicts a pseudo-clear image and we perform an adaptive fusion to obtain the final result. Note that the global density features in Global Branch are fed into Local Branch by cross-branch connections.

Overall, the main contributions of our work are as follows:

- We propose to learn and refine the haze density feature of a hazy image by utilizing the difference information between a generated proposal image ($\mathbf{P}$) and a hazy input ($\mathbf{I}$) and an end-to-end method named DFR-Net is designed to achieve density-aware dehazing.
- We extract haze density features both globally and locally. In GB, a Global Block is introduced, which can explore the image features of $\mathbf{P}$ and $\mathbf{I}$. To highlight the features that can better describe the density information, a GDFR module is proposed. In LB, local density features are extracted from the dehazing residual between $\mathbf{P}$ and $\mathbf{I}$. Additionally, a IDRF module is presented to refine local density features stage by stage.
- Sufficient experiments are conducted on our DFR-Net and demonstrate that it can achieve better results over the existing state-of-the-art (SOTA) methods on multiple commonly used datasets.

II. RELATED WORKS

A. Density-aware Dehazing Methods

In recent years, several methods [15–18, 23–25] have attempted to improve the dehazing performance by enabling the network to perceive haze density.

1) Density-awareness via estimating T-map: Haze density is influenced by several factors and is inversely proportional to T-map, so some methods learn density information by estimating T-map. Lou et al. [19] predict a T-map first for nighttime image dehazing. Zhang et al. [16] estimate a low-resolution T-map and then jointly input the feature map and the estimated T-map to a Laplacian pyramid decoder to achieve a restored image. Yang et al. [18] propose a semi-supervised method that does not require paired data. The method estimates T-map, scattering coefficient, and depth to reconstruction hazy images and restores clear images. However, these methods require additional labeled data and might be inaccurate due to the complexity of practical scenes [12].
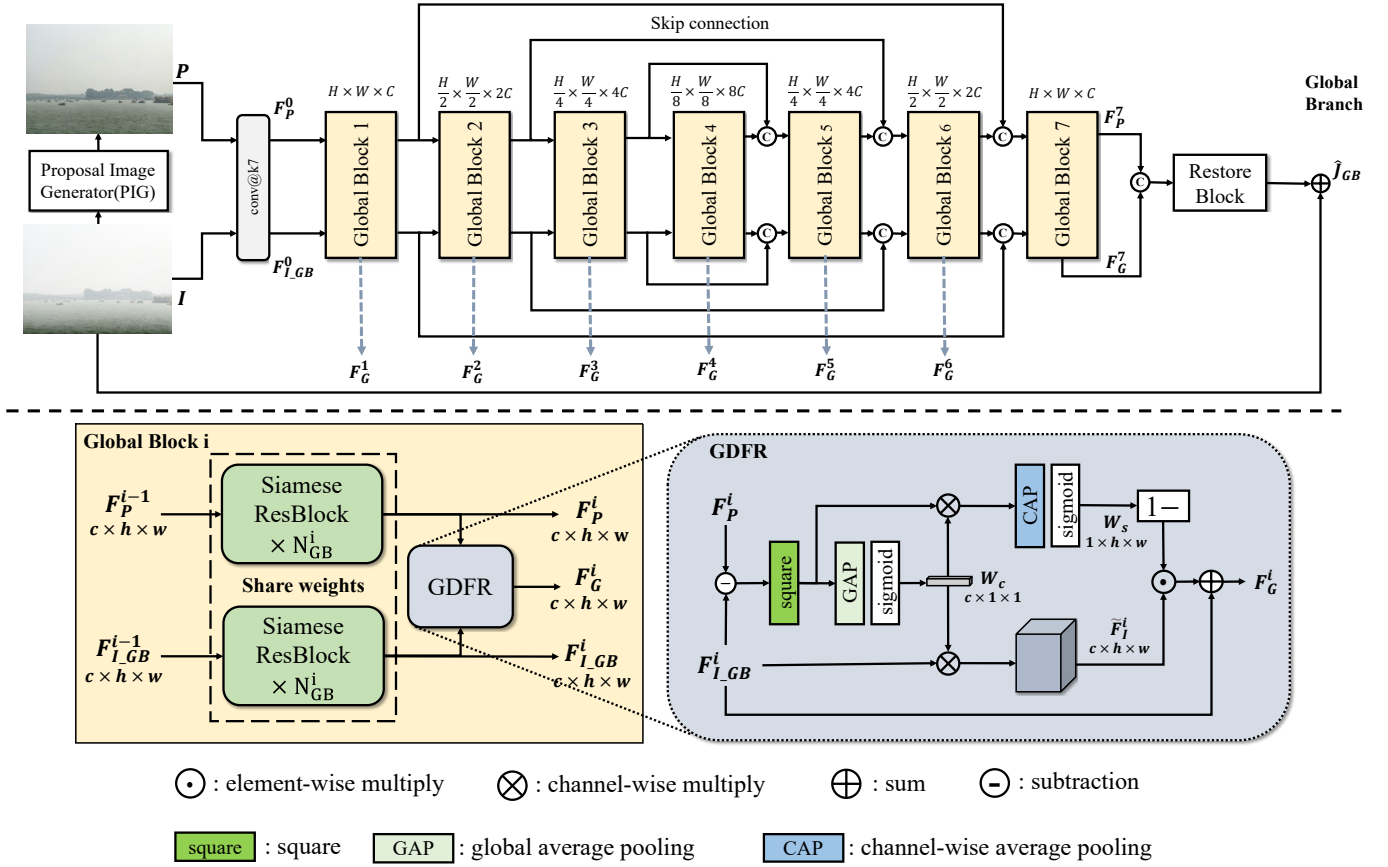
Fig. 3. The illustration of GB, global block and GDFR module. Note the global density features are fed into LB by cross-branch connections.

2) Density-awareness via extracting density features directly: Research works [15, 17, 26] directly learn haze density information without estimating a T-map. Deng et al. [17] design a Haze-Aware Representation Distillation (HARD) module to extract global brightness and a haze-aware map. Chen et al. [26] propose an attention mechanism based on dark channel prior to describe haze concentration. However, not estimating the T-map would result in a lack of a comparator to measure density. Generating intermediate results and using the information contained therein can address this issue.

B. Dehazing Methods with Intermediate Results

Considering the difficulty of recovering images directly from the haze input, dehazing methods [15, 20, 21, 27] which generate intermediate results (or one result) inside the network to facilitate the dehazing process are proposed. Bai et al. [20] first generate a reference image by a deep pre-dehazer, and then develop a progressive feature fusion module to fuse the hazy and reference features, which achieves high metrics on several datasets. Chen et al. [21] first remove light and thick smoke by a Smoke Remove Network (SRN) to gain a coarse output, which is concatenated with the original input and fed to a Pixel Compensation Network (PCN) to recover the missing pixels in the thick smoke. Hong et al. [27] propose an Uncertainty-Driven Dehazing Network. In this method, intermediate results are together generated with uncertainty maps for uncertainty features extraction. Ye et al. [15] also

pre-generate a pseudo-haze-free image. The hazy input and the pseudo-haze-free image are concatenated to estimate a Density Encoding Matrix describing the relationship between haze density and absolute position and mixed up to the following deep layers.

Despite the above methods extracting feature from intermediate results, they do not fully consider the differences between these results and the haze inputs, especially the differences in haze density. Simple concatenation [20, 21] or linear summation [15] might lead the networks to rely on the uncertain learning process and lose the capture of information about the differences between the two images. In addition, the lack of a targeted design that addresses the relationship between the intermediate results and the original input leads the extracted features not fine enough and limits the dehazing performance.

Our DFR-Net improves on the aforementioned methods by exploring and refining density features through the utilization of density differences between a generated proposal image and the hazy input, thereby achieving an awareness of haze density and superior dehazing performance.

III. METHOD

As illustrated in Fig. 2, DFR-Net generates a proposal image using a Proposal Image Generator (PIG), which is a simple U-Net, to facilitate density-awareness. Besides, DFR-Net consists of two primary parts: Global Branch (GB) and
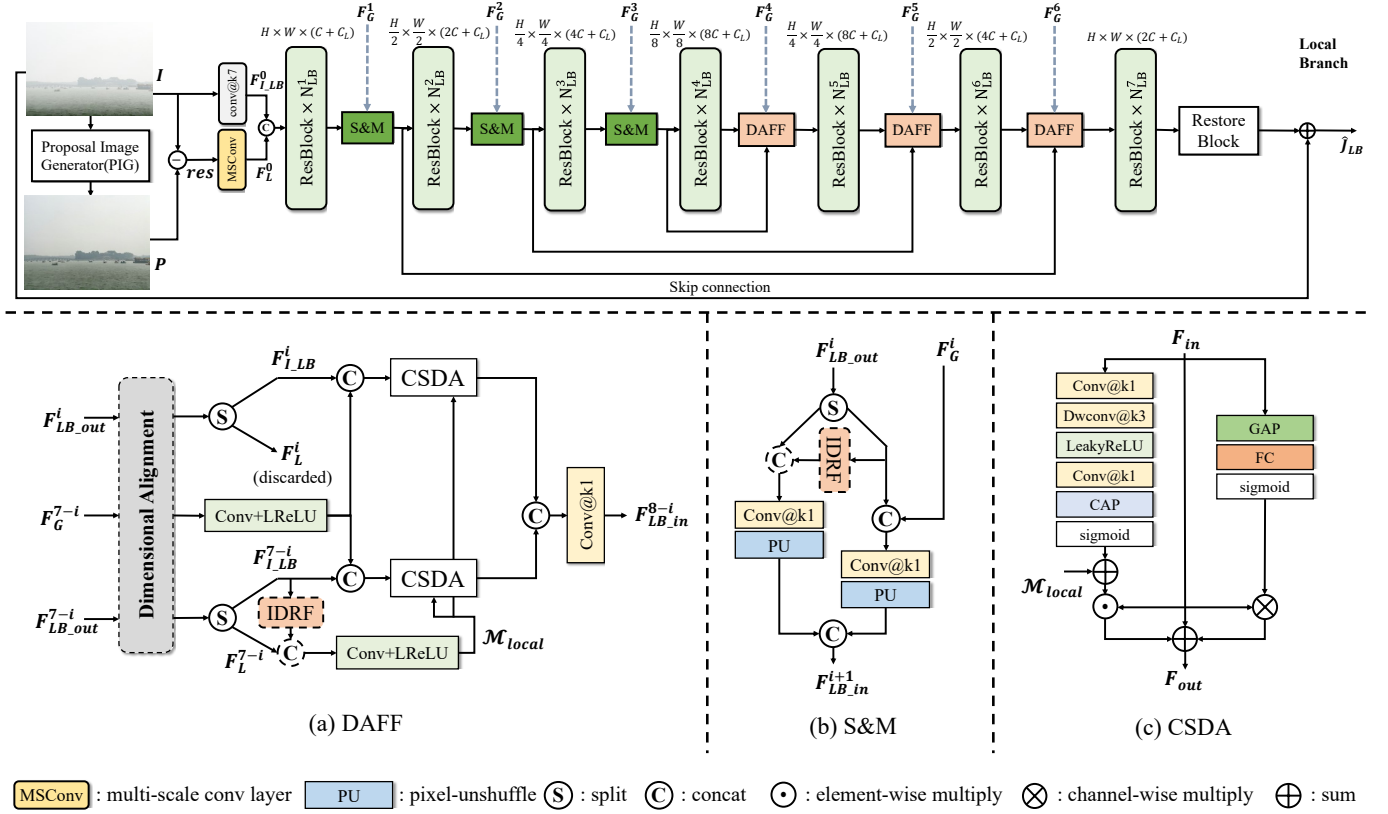
Fig. 4. The illustration of the structure of LB (top) and DAFF (a) , S&M (b), CSDA (c) modules. The locations of the IDRF usage are indicated by dashed lines representing plug-and-play availability.

Local Branch (LB). The GB and LB are responsible for learning and refining global and local haze density features, respectively, and generate a pseudo result each. In the end, an adaptive fusion is used to obtain the final dehazing result. This section presents a detailed introduction of our method's main ideas.

A. Proposal Image for Density-awareness

In DFR-Net, we first generate a proposal image (**P**) that provides information about the difference in haze density. **P** exhibits lower overall density than the input image (**I**), and inconsistent dehazing performance in some local areas. Taking these characteristics into account, we are motivated to extract and refine haze density features in the subsequent network.

To generate **P**, we employ a Proposal Image Generator (PIG), which is a simple U-net comprising multiple Residual Blocks (ResBlocks). We pre-train PIG with paired images and incorporate it into subsequent branches as an end-to-end network.

B. Global Branch

1) Overview: The Global Branch (GB) is designed to extract and refine global density features using the overall haze density difference. As depicted in Fig. 3 (top), GB consists of a 7-stage U-Net with 7 global blocks, where each block has a Siamese structure with $N_{GB}^i (i \in [1, ..., 7])$ ResBlocks for feature extraction and a GDFR module for

global density feature refinement. The inputs to GB are **I** and **P** $\in \mathbb{R}^{H \times W \times 3}$, which are embedded to $F_{I\_GB}^0$, $F_P^0 \in \mathbb{R}^{H \times W \times C}$ using a convolutional layer. Upsampling or downsampling is performed between every two blocks. To obtain the predicted dehazing residual, the output features of the last global block are fed to a restore block, and the predicted dehazing residual is added to **I** to obtain a pseudo result, $\hat{J}_{GB}$.

We design the global block as a basic unit in GB, taking the haze density relationship between **I** and **P** into consideration. Specifically, the i-th global block is composed of a Siamese structure with N_{GB}^i ResBlocks for feature extraction and a GDFR module for global density feature refinement. For the i-th global block, the inputs are the outputs of the previous block, $F_{I\_GB}^{i-1}$ and F_P^{i-1}, the outputs are $F_{I\_GB}^i$, F_P^i and the refined global density feature, F_G^i.

2) Siamese ResBlocks for global feature extraction: Siamese structure shares weights and can measure similarity [28–30] or dissimilarity [31, 32] between samples effectively. So we utilize a Siamese structure to extract image features from both **P** and **I**, which enables the network to establish a relationship between images with different haze densities. Moreover, compared to a single hazy input structure, the Siamese structure enables the features extracted by the network to better perceive the variation in haze density.

3) GDFR for global feature refinement: To further refine the extracted features, we propose the Global Density Feature Refinement (GDFR) module, which aims to highlight the features that better describe global density information and

thus pull apart the features of images with different densities. For the implementation, we compute the difference between the feature maps of **I** and **P**, and then square each element in the difference map. This subtractive operation filters out non-density information, and the resulting feature channels with larger differences can better describe density information. We perform global average pooling (GAP) and sigmoid operation on the feature differences to obtain density-related channel weights, denoted as W_c. We then multiply these weights with $F_{I\_GB}^i$ channel-wise:

$$\tilde{F}_I^i = F_{I\_GB}^i \otimes \sigma(GAP(POW((F_P^i - F_{I\_GB}^i), 2))) \quad (2)$$

where $GAP(\cdot)$ stands for global average pooling, $\sigma(\cdot)$ represents sigmoid operation, $\otimes$ stands for channel-wise multiply and $POW(a, b)$ represents a power of exponent b for each element of a.

In spatial dimension, we first multiply the channel weights by the squared feature difference. Next, we apply channel-wise average pooling (CAP) and sigmoid operations to obtain a 2-D weight map that represents the spatial density difference W_s. Higher value in W_s indicates that haze has been more effectively removed here. Therefore, $1 - W_s$ can represent an attention map that guides the network's focus. Here 1 denotes a 2-D tensor of ones with the same shape of W_s. We then multiply $1 - W_s$ and $\tilde{F}_I^i$ and add the result to $F_{I\_GB}^i$ to obtain the finally refined global density feature F_G^i:

$$F_G^i = (1 - \sigma(CAP(POW((F_P^i - F_{I\_GB}^i), 2)))) \odot \tilde{F}_I^i + F_{I\_GB}^i \tag{3}$$

where $CAP(\cdot)$ stands for channel-wise average pooling.

After the last block, we concatenate F_P^7 and F_G^7 and input the concatenated feature to a restore block, which is composed of four ResBlocks and a convolutional layer, to obtain the dehazing residual.

$$res_{GB} = RB_{GB}(cat(F_P^7, F_G^7)) \tag{4}$$

where $RB_{GB}(\cdot)$ denotes the calculation of the restore block in GB and $cat(\cdot)$ denotes the channel-wise concatenation operation. And the pseudo-clear image of GB can be obtained by: $\hat{J}_{GB} = I + res_{GB}$.

C. Local Branch

1) Overview: LB extracts local density features from the dehazing residual of **P** and **I** , which contains local density differences and indicates hard dehazing or high haze density areas. To refine these local features, we propose the IDRF module that gradually adjusts the features to match those of clear images.

The main body of LB is a 7-stage U-net, as illustrated in Fig. 4 (top). Each stage is composed of $N_{LB}^i(i \in [1,...,7])$ ResBlocks. LB takes **I** and **res** $\in \mathbb{R}^{H \times W \times 3}$ as inputs and embeds them to shallow features $F_{I\_LB}^0 \in \mathbb{R}^{H \times W \times C}$, $F_L^0 \in \mathbb{R}^{H \times W \times C_L}$ using a normal convolutional layer and a multi-scale convolutional layer, respectively. In the encoder stages, local and image features are extracted by a split and merge way, and in the decoder stages, DAFF fuses shallow image features, deep image features, local features, and refined global

features from GB. The dehazing residual can be obtained by a restore block, and the pseudo result $\hat{J}_{LB}$ can be obtained by summing it and **I**.

2) S&M for local feature extraction: The local features are extracted in an S&M way. Firstly, $F_{I\_LB}^0$ and F_L^0 are merged by concatenation and fed into stage-1, resulting in an output feature with $(C + C_L)$ channels. The output feature is then processed by the S&M module. As shown in Fig. 4 (b), the input feature is split into an image feature and a local density feature. Notably, the number of channels for the image feature varies from stage to stage, while the number of channels for the local feature is consistently fixed at C_L. To utilize the global density information obtained from GB, the image feature is concatenated with the refined global feature from GB. Afterwards, the concatenated feature and local feature are downsampled by 1×1 convolutional layers and pixel-unshuffle, respectively. The downscaled features are then concatenated and fed into the next stage. This design allows local features to fully interact with hazy image features, thus enhancing the reliability of local features.

3) DAFF for feature fusion: Several methods introduce skip connection to aggregate shallow and deep features and simply concatenate or add them together [33, 34], which might result in a loss of information. To fully utilize the information relevant to density, we employ Density Aware Feature Fusion (DAFF). Specifically, given a set of features: $\{F_{LB\_out}^i, F_{LB\_out}^{7-i}, F_G^{7-i}\}(i \in [1, 2, 3])$, which represent the output feature of the i-th, $(7-i)$-th stage of LB and the refined global feature of the $(7-i)$-th block in GB, we first align them to a same shape and split the image features and local density features. Then we perform convolutions and LeakyReLU (Conv+LReLU) on F_G^{7-i} and the results are concatenated with $F_{I\_LB}^i$ and $F_{I\_LB}^{7-i}$ respectively and input to Channel-Spatial Density Attention (CSDA) modules, as illustrated in Fig. 4 (c). Moreover, the local feature F_L^{7-i} is projected to a 2-D local attention map ($\mathcal{M}_{local}$) by Conv+LReLu and $\mathcal{M}_{local}$ is fed to CSDAs. Finally, the outputs of CSDAs are concatenated and compressed to the channel number of the subsequent stage by a 1×1 convolutional layer.

4) IDRF for local feature refinement: To refine the local density feature, we further introduce IDRF module. As shown in Fig. 5, IDRF takes $F_{I\_LB}^i$ as input to obtain intermediate dehazing residual res_{inter}^i by an intermediate restore block (IRB) composed of two ResBlocks and a convolutional layer. Then the res_{inter}^i is projected to a C_L-channel feature embedding $F_L^{'}$ which is subsequently concatenated with local features as shown in Fig. 4 (a) and (b). With this process, local density features can be updated by the current dehazing residual. To optimize this module and pull local features in towards clear image features, we employ a local density refinement loss ($\mathcal{L}_{LDR}$) which will be introduced in Sec. III-D.

Similar to GB, LB predicts a pseudo-clear result: $\hat{J}_{LB} = I + res_{LB}$. Finally we fuse the pseudo-results of the two branches with a learnable parameter α: $\hat{J} = \alpha \times \hat{J}_{GB} + (1 - \alpha) \times \hat{J}_{LB}$.

D. Loss Function

The overall loss fuction of our DFR-Net can be formulated as: $\mathcal{L} = \mathcal{L}_{Rec} + \lambda_1 \mathcal{L}_P + \lambda_2 \mathcal{L}_{RD} + \lambda_3 \mathcal{L}_{LDR}$, where $\mathcal{L}_{Rec}$ and

TABLE I
QUANTITATIVE COMPARISON OF DFR-NET WITH THE STATE-OF-THE-ART IMAGE DEHAZING METHODS ON DIFFERENT DATASETS (PSNR (DB)/SSIM).
BEST RESULTS ARE **BOLDED** AND SECOND BEST RESULTS ARE <u>UNDERLINED</u>. CELLS WHERE RESULTS ARE NOT AVAILABLE ARE REPLACED BY "-"

| Method	RESIDE-outdoor		Haze4K		NH-HAZE		Dense-Haze	
	PSNR(dB)	SSIM	PSNR(dB)	SSIM	PSNR(dB)	SSIM	PSNR(dB)	SSIM
Non-Density-Aware								
DCP [6] (TPAMI' 10)	19.13	0.815	14.01	0.760	10.57	0.522	11.01	0.416
DehazeNet [35] (TIP' 16)	24.75	0.927	19.12	0.840	12.86	0.545	9.48	0.438
AOD-Net [12] (ICCV' 17)	24.14	0.920	17.15	0.830	15.40	0.571	12.82	0.468
GDNet [36] (ICCV' 19)	30.86	0.982	23.29	0.930	18.33	0.667	14.96	0.530
MSBDN [37] (CVPR' 20)	33.48	0.982	22.99	0.850	19.23	0.713	15.13	0.555
FFA-Net [38] (AAAI' 20)	33.57	0.984	26.96	0.950	19.87	0.694	12.22	0.444
AECR-Net [14] (CVPR' 21)	-	-	-	-	19.88	0.722	15.80	0.466
CEEF [3] (TMM' 22)	19.13	0.792	-	-	-	-	-	-
SGID-PFF [20] (TIP' 22)	30.20	0.975	-	-	-	-	-	-
UDN [27] (AAAI' 22)	34.92	0.987	-	-	-	-	-	-
QCNN-H [39] (TC' 23)	28.74	0.964	-	-	-	-	-	-
MFINEA [40] (NN' 23)	33.88	0.981	-	-	-	-	<u>18.34</u>	<u>0.609</u>
DehazeFormer-B [41] (TIP' 23)	34.95	0.984	30.29	<u>0.985</u>	17.37	0.725	-	-
Density-Aware								
HDDNet [16] (TC' 22)	22.52	0.910	-	-	-	-	-	-
DeHamer [23] (CVPR' 22)	<u>35.18</u>	0.986	-	-	<u>20.66</u>	0.684	16.62	0.560
PMNet [15] (ECCV' 22)	34.74	<u>0.990</u>	<u>33.49</u>	0.980	20.42	<u>0.731</u>	16.79	0.510
DFR-Net (ours)	**35.34**	**0.993**	**34.63**	**0.993**	**21.21**	**0.810**	**18.85**	**0.674**
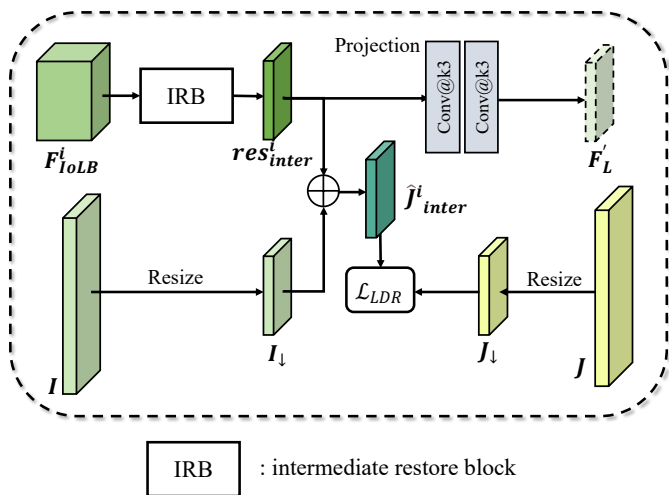
Fig. 5. The illustration of IDRF module. The projected feature F_L' will be used to update local features in corresponding S&M or DAFF module.

IRB : intermediate restore block

$\mathcal{L}_P$ denote L1 loss and perceptual loss [44] between the predicted haze-free output $\hat{J}$ and ground truth J, $\mathcal{L}_{RD}$ and $\mathcal{L}_{LDR}$ represent representation dissimilarity loss and local density refinement loss, and λ_1, λ_2 and λ_3 are hyper-parameters for loss regulation.

1) Representation Dissimilarity Loss: In this paper, we introduce a Siamese structure to learn density-related features from **I** and **P**. To motivate the Siamese structure to learn more information about the differences between the inputs, we design representation dissimilarity loss:

$$\mathcal{L}_{RD} = \sum_{i=1}^{n} \langle F_P^i, F_{I\_GB}^i \rangle \qquad (5)$$

where $\langle a, b \rangle$ represents the calculation of the cosine similarity between a and b. We compute cosine similarities between the intermediate features of **I** and **P** and minimize them to let the Siamese structure learn more representation about the difference between the two inputs.

2) Local Density Refinement Loss: To achieve the pulling of local features from haze images to clear images, we introduce local density refinement loss.

$$\mathcal{L}_{LDR} = \frac{1}{k} \sum_{j=1}^{k} \|\hat{J}_{inter}^j(x) - Down^j(J(x))\|_1 \qquad (6)$$

where k denotes the number of IDRF modules which are applied, $\hat{J}_{inter}^j(x)$ is the j-th intermediate predicted clear output, and $Down^j(\cdot)$ is the operation that downsamples ground truth to the size of the corresponding intermediate output.

IV. EXPERIMENTS

A. Datasets and Metrics

We conduct experiments on several datasets to train our method and test our method's dehazing performance. The datasets include: (1) RESIDE outdoor [45], which contains 313950 synthetic outdoor hazy/clear image pairs for training and 500 pairs for testing; (2) Haze4K [42], which includes a training set of 3000 indoor-outdoor mixed image pairs and a testing set of 1000 image pairs; (3) NH-HAZE [46], a real-world dataset for the NTIRE 2020 competition, which consists of 55 pairs of non-homogeneous hazy images and clear images of real scenes (45, 5 and 5 pairs for training, validation and test respectively); (4) Dense-Haze [43, 47], a real-world dataset for the NTIRE 2019 competition and contains 55 pairs of dense-haze images and corresponding clear images (with same data split as NH-HAZE). We evaluate the dehazing effectiveness of our method using two commonly used image quality metrics: PSNR (dB) and SSIM.
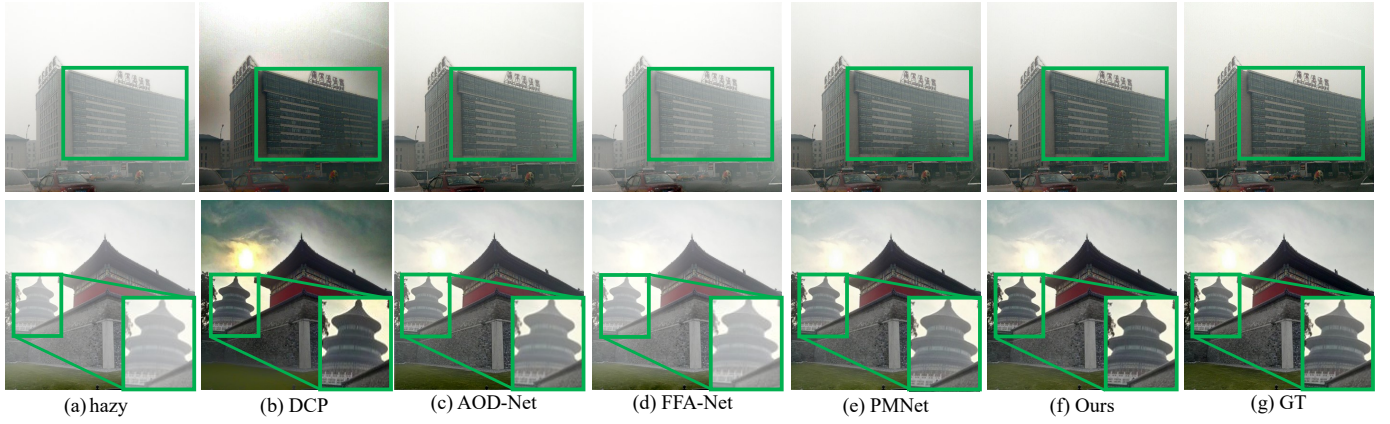
Fig. 6. Visual comparison of various methods on Haze4K [42] dataset. Areas where our method works better are boxed out and zoomed in, or you can zoom in yourself to get a better view.
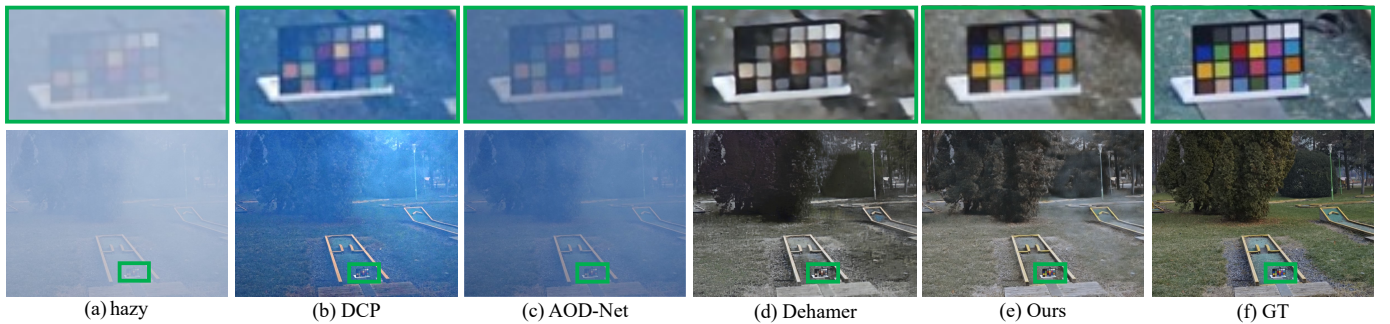

Fig. 7. Visual comparison of various methods on Dense-Haze [43] dataset. Areas where our method works better are boxed out and zoomed in, or you can zoom in yourself to get a better view.

B. Implementation Details

1) Network Configuration: In our work, we use the Basic Block structure proposed in [38] as our ResBlock. $\{N_{GB}^1, N_{GB}^2, N_{GB}^3, N_{GB}^4, N_{GB}^5, N_{GB}^6, N_{GB}^7\}$ and $\{N_{LB}^1, N_{LB}^2, N_{LB}^3, N_{LB}^4, N_{LB}^5, N_{LB}^6, N_{LB}^7\}$ are set to $\{2, 2, 3, 4, 3, 2, 2\}$ and $\{4, 6, 8, 10, 6, 8, 8\}$. In particular, we set the basic feature dimension of GB and LB, C, to 32 and the local density dimension C_L to 4. All upsampling or downsampling operations are implemented by 1×1 convolution with pixel-shuffle or pixel-unshuffle. And we use IDRF module after every stage in LB except the last stage.

2) Training Settings: The PIG is pre-trained and spliced with GB and LB. And the whole network is trained in an end-to-end fashion. We use AdamW optimizer ($\beta_1 = 0.9$, $\beta_2 = 0.999$, weight decay is $1e^{-4}$) to train the model and iterate 600k times with the initial learning rate $1e^{-4}$ reduced to $1e^{-6}$ with the cosine annealing [48]. Following [33], we perform progressive learning in our training process, which leads the network to adapt to those inputs close to the size in practical applications. For loss regulation, we set $\lambda_1 = 0.2$, $\lambda_2 = 0.001$ and $\lambda_3 = 0.1$. The data augmentations include random cropping, horizontal flipping, and vertical flipping.

C. Comparison with State-of-the-art Methods

Our comparison methods include DCP [6], DehazeNet [35], AOD-Net [12], GDNet [36], MSBDN [37], FFA-Net [38], AECR-Net [14], CEEF [3], SGID-PFF [20], UDN [27], QCNN-H [39], MFINEA [40] , DehazeFormer [41], HDDNet [16], DeHamer[23] and PMNet [15].

1) Quantitative Evaluations: The quantitative results are shown in Table I. It demonstrates that our method achieves the highest metrics on RESIDE-outdoor, Haze4K, Dense-Haze and NH-HAZE datasets, outperforming other SOTA methods. Notably, DFR-Net achieves a 1.14dB PSNR gain on the Haze4K dataset and significant improvements in SSIM on all datasets.

Among the density-aware methods, our DFR-Net outperforms all other methods [15–17, 23]. Compared to the transmission-aware (density-related) methods DeHamer [23] and HDDNet [16], we directly extract density features from hazy images and obtain better performance. DFR-Net also surpasses PMNet on metrics on multiple datasets by utilizing density difference information between **I** and **P**, rather than simply concatenating them.

2) Qualitive Evaluations: Fig. 6 shows visual comparisons between our DFR-Net and SOTA methods on Haze4K dataset, demonstrating that our DFR-Net can more effectively remove haze than other methods. Specifically, AOD-Net [12] and FFA-Net [38] still leave haze residue in most areas, DCP [6] removes some of the haze but suffers from color distorion. While PMNet [15] achieves better dehazing performance than previous methods, our method is able to recover clearer
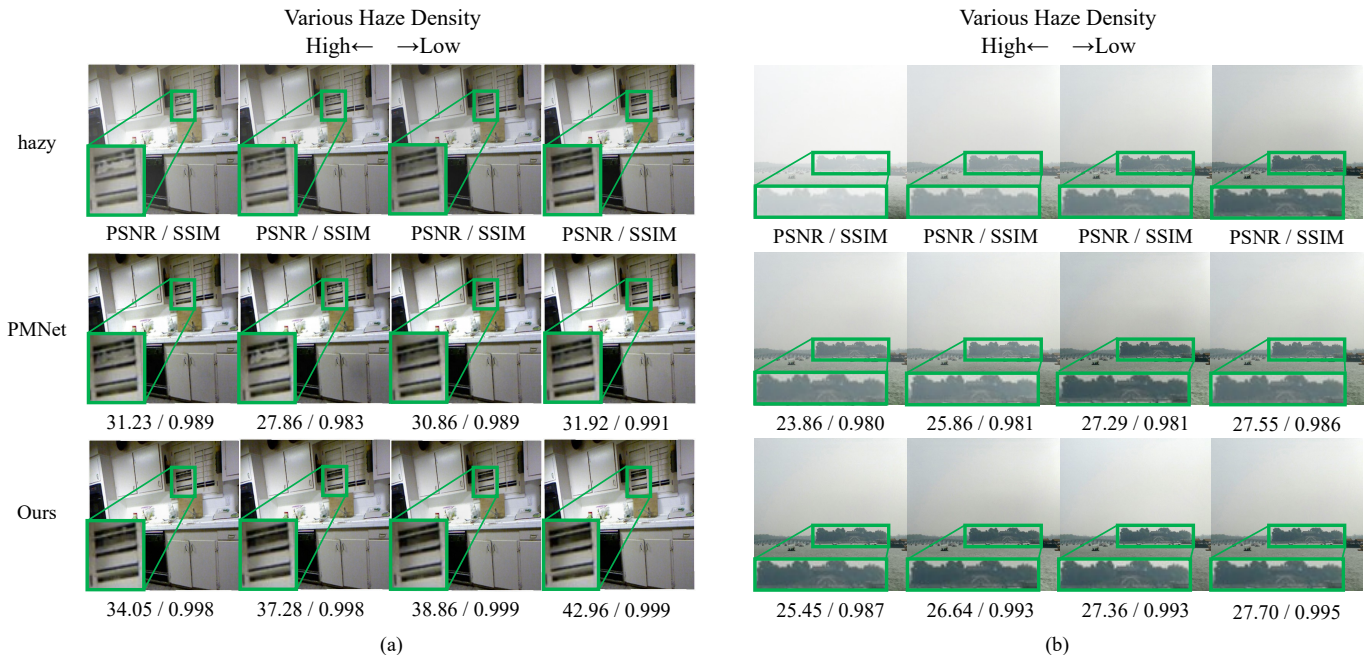
Fig. 8. Visual comparison with PMNet on Haze4K [42] dataset. Note that the densities of the input images are varied for each set of images of the same scene, gradually decreasing from left to right. Please zoom in for a better view.

images.

In addition, we compare the visual quality performance of our method with other methods on real-world dataset (Dense-Haze) in Fig. 7. Compared to other methods, our DFR-Net removes the overall haze while achieving fine restoration in image details. This is attributed to the pulling in of the image features towards a clear image at each stage in LB, which reduces the occurrence of color distortion and blurring of details in our results.

We further compare the visual results of our DFR-Net with those of PMNet [15] on the Haze4K dataset. Fig. 8 illustrates the dehazing outcomes of PMNet and DFR-Net for both indoor and outdoor images with varying haze densities. DFR-Net exhibits more consistent dehazing performance across different haze densities compared to PMNet. In particular, DFR-Net demonstrates superior performance in preserving local image details, as depicted in Fig. 8 (a). Conversely, PMNet struggles to accurately restore local details in the dehazed images. Additionally, DFR-Net outperforms PMNet in handling images with high global haze density, as shown in Fig. 8 (b). The incorporation of both global and local density difference information in DFR-Net contributes to its robustness in perceiving and effectively addressing different haze densities. The global density component enables the network to better comprehend variations in haze densities, allowing for improved performance on images with diverse densities. On the other hand, the local density component motivates the network to identify and restore details in regions with high density, resulting in finer dehazed images with enhanced visual quality. Further details and discussions regarding these aspects are presented in Sec. IV-D.

D. Ablation Studies

The main innovation of DFR-Net is the extraction and utilization of global and local haze density information. Therefore, we conduct ablation experiments and analyses in this section to demonstrate the effectiveness of our utilization of haze density information.

We conduct ablation experiments on the modules and loss functions proposed in this paper to evaluate their effectiveness on dehazing. We first establish a **base** network (①), which consists of a non-weight sharing GB (pseudo-Siamese structure), and an LB that takes the directly concatenated **I** and **P** as input and aggregates features by concatenation. Only $\mathcal{L}_{Rec}$ and $\mathcal{L}_P$ are employed to optimize the **base** network. Then we define several variants to verify the effectiveness of our proposed modules and loss functions on dehazing: ② **+Siamese**: Use the Siamese structure for feature extraction. ③ **+$\mathcal{L}_{RD}$**: Incorporate $\mathcal{L}_{RD}$ into the total loss function. ④ **+GDFR**: Use GDFR to obtain refined global density features. ⑤ **+DAFF**: Aggregate features by DAFF. ⑥ **+DR**: Explore local density from **D**ehazing **R**esidual (DR) in LB. ⑦ **+IDRF**: Use IDRF after each stage of LB, except the last one. ⑧ **+$\mathcal{L}_{LDR}$** (our default setting): Include $\mathcal{L}_{LDR}$ in the total loss function. Among them, variants ②-④ and ⑤-⑧ gradually introduce global and local haze density respectively. The ablation results are presented in Table II. Subsequently, we will analyze the effectiveness of introducing global and local haze density separately.

1) Ablation studies on utilizing global density difference and global feature refinement: We carry out experiments on variants ②, ③, ④ to verify the effectiveness of our design using global density difference to extract and refine global density-related features. To ensure fairness, we adjust the

TABLE II
ABLATION STUDIES ON PROPOSED MODULES AND LOSS FUNCTIONS ON THE HAZE4K DATASET. NOTE THAT FLOPs AND PARAMS ARE MEASURED ON 256×256 IMAGES.

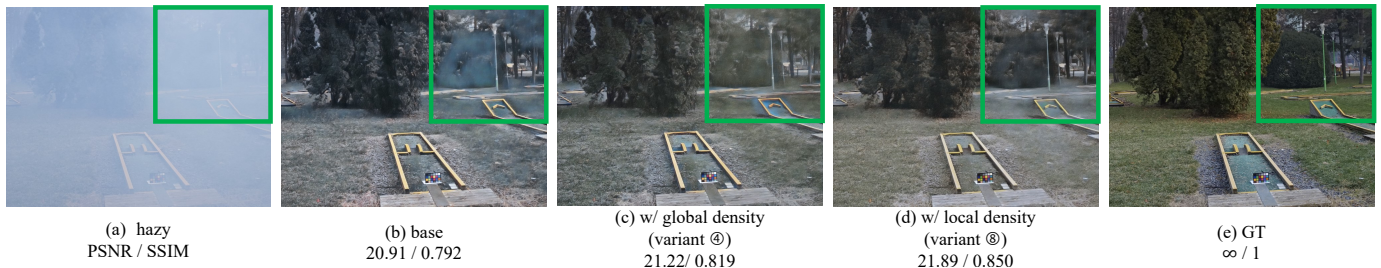|	Label	Setting	PSNR (dB)	FLOPs (G)	Params (M)
w/o density difference	①	base	30.52	222.46	35.76
+ global density difference	②	+Siamese	31.21	241.94	35.15
	③	$+\mathcal{L}_{RD}$	31.55	241.94	35.15
	④	+GDFR	32.65	242.63	35.15
+ local density difference	⑤	+DAFF	32.89	271.54	40.52
	⑥	+DR	33.01	271.56	40.52
	⑦	+IDRF	33.52	286.60	42.11
	⑧	$+\mathcal{L}_{LDR}$ (default)	**34.63**	286.60	42.11

Fig. 9. Visualization of the ablation experiments on exploring and utilizing global / local haze density difference. Areas with large variance in dehazing effectiveness are framed out. Please zoom in for a better view.

(a) hazy
PSNR / SSIM

(b) base
20.91 / 0.792

(c) w/ global density
(variant ④)
21.22 / 0.819

(d) w/ local density
(variant ⑧)
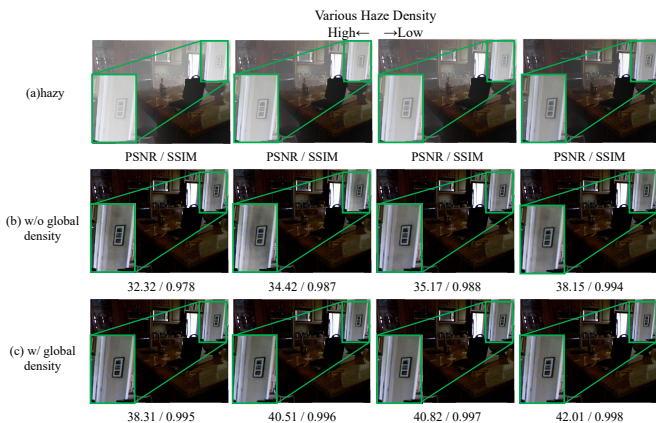21.89 / 0.850

(e) GT
∞ / 1

Fig. 10. Visualization of the ablation experiments on utilizing global density difference and global feature refinement. Note that the haze densities of the input image are varied and gradually decrease from left to right. Please zoom in for a better view.
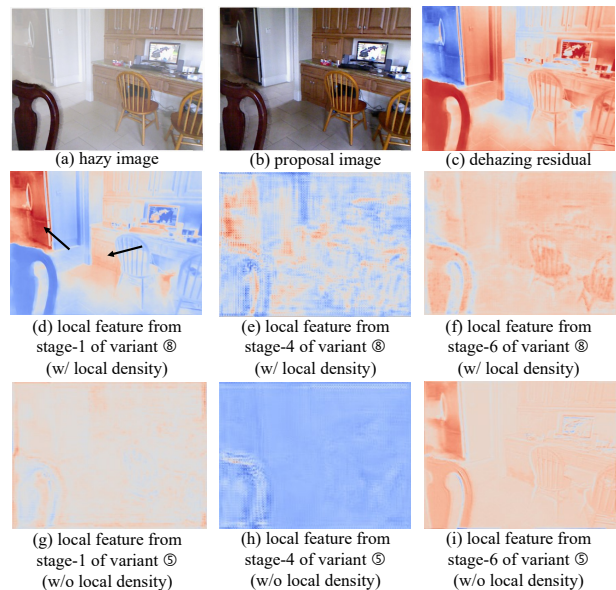

Fig. 11. Visualization (heatmap form) of the local density features from variant ⑧ and ⑤. The black arrows in (d) point out the hard regions found by the network by utilizing the dehazing residual.

ResBlock numbers of the Siamese structure to approximately match the number of parameters in the base network. Table II indicates that sharing parameters enables better feature learning from proposal images and delivers a 0.69 dB PSNR improvement. To motivate the network to learn more density-related information, $\mathcal{L}_{RD}$ is employed and the inclusion of it leads to a 0.34 dB PSNR improvement. In addition, the introduction of GDFR brings 1.10 dB PSNR performance gain by refining global densities features.

We also conducted experiments to validate the effectiveness of incorporating global density difference information. Fig. 10 showcases the dehazing results obtained without (variant ①)

and with (variant ④) global density information, using image inputs with varying densities. When global density difference information is not introduced, we observed low quantitative metrics, unstable dehazing results, and inconsistent performance within the same scene. As depicted in Figure 10, variant ① produces suboptimal dehazing outcomes.

However, upon incorporating global density differences, the network exhibits improved capability to perceive variations in

TABLE III
ABLATION STUDIES ON THE USAGE OF IDRF. ✔ INDICATES THAT THE
OUTPUT FEATURES OF THIS STAGE WILL BE PROCESSED FOR IDRF.

	Stage						PSNR (dB)	Params (M)
	1	2	3	4	5	6		
IDRF	✔	✔	✔				33.06	41.32
				✔	✔	✔	32.89	41.32
	✔	✔	✔	✔	✔	✔	**33.52**	42.11

TABLE IV
ABLATION STUDIES ON THE EFFECTIVENESS OF DAFF. THE
EXPERIMENTS ARE CONDUCTED ON THE HAZE4K DATASET.

Methods	concat [15]	SK Fusion [41]	DAFF
PSNR (dB)	33.19	33.80	**34.63**

global densities. This is achieved through the exploration of global density differences between the proposal image and the input image using the Siamese structure, as well as the refinement of density features using the Global Density Feature Refinement (GDFR) module. As a result, the network becomes more robust to images with different density levels, enabling the generation of consistent and visually clear dehazed images.

2) Ablation studies on utilizing local density difference and local feature refinement: We perform experiments on variants ⑥, ⑦, ⑧ for validation of the effectiveness of utilizing local density difference. As shown in Table II, extracting local features from dehazing residual, IDRF and $\mathcal{L}_{LDR}$ result in 0.12, 0.41 and 1.11 dB PSNR performance gains.

To demonstrate the necessity of ultilizing local density difference, intermediate local features of the network using dehazing residual to extract features (variant ⑧) or not (variant ⑤) are visualized in Fig. 11, which shows that high density and hard dehazing areas are captured in the shallow stage of our method. And as the network deepens, the local maps are gradually flattened, which implies the local features are updated stage by stage. However, variant ⑤ fails to achieve these performances.

3) Ablation studies on the location of IDRF usage: We carried out ablation experiments to determine the optimal placement of the IDRF module within the network architecture. We compared three cases: using IDRF only in the encoder stages, only in the decoder stages, and in all stages. The results, as shown in Table IV-D1, indicate that employing IDRF in all stages yields the best overall performance.

Furthermore, we observed that incorporating IDRF in the encoder stage leads to improved quantitative metrics compared to its use in the decoder stages. This can be attributed to the fact that the refinement of local features should commence as early as possible in the network. The primary role of IDRF is to update the current local features and facilitate the alignment of the restored image features with clear image features. Without this refinement process, the network's local features remain relatively unchanged, resulting in less attention being paid to regions with relatively high density or hard features that require careful handling at the current stage. By introducing IDRF at an earlier stage, the network can be guided more effectively to bring the image features closer to the features of clear images.

4) Ablation studies on the effectiveness of DAFF: We further conduct ablation experiments on the effectiveness of DAFF. DAFF is designed to fully fuse the global density features passed by cross-branch connections, the image features of the LB branch, and the local density features. And

we compare the PSNR performance with other two fusion strategies: concatenation [15] and SK Fusion [41]. Table IV demonstrate that our DAFF achieves better results than the other two methods.

As described in Section III-C3, DAFF first combines the global features with the image features, enabling the features to be aware of global density information. It then introduces the local features through the CSDA mechanism, facilitating pixel-level refinement of the features. In contrast, both concatenation and SK Fusion methods focus primarily on channel attention, without considering the attentional role of the local density map on the spatial dimensions of the features. The superior performance of DAFF can be attributed to its comprehensive fusion strategy, which takes into account both global and local density information. This enables our method to capture and utilize density-related information more effectively, resulting in improved dehazing performance compared to concatenation and SK Fusion methods.

V. CONCLUSION

In this paper, we propose DFR-Net, a density-aware method for image dehazing that utilizes haze density differences to extract and refine density-related features. To achieve this, we generate a proposal image and explore density representation from it and the hazy input. We use two branches to extract and refine global and local density features, respectively, based on the density differences between the proposal image and the hazy input. In Global Branch, features of images with different densities are pushed away, while in Local Branch, hazy image features are gradually pulled closer to those of clear images. Our experimental results demonstrate that DFR-Net achieves high-performance image dehazing ability, and our ablation studies show that our designs enable fine awareness and refinement of density information.

REFERENCES

[1] Cunyi Lin, Xianwei Rong, and Xiaoyan Yu. Msaff-net: Multiscale attention feature fusion networks for single image dehazing and beyond. *IEEE transactions on multimedia*, 2022. 1

[2] Chongyi Li, Chunle Guo, Jichang Guo, Ping Han, Huazhu Fu, and Runmin Cong. Pdr-net: Perception-inspired single image dehazing network with refinement. *IEEE Transactions on Multimedia*, 22(3):704–716, 2019.

[3] Xiaoning Liu, Hui Li, and Ce Zhu. Joint contrast enhancement and exposure fusion for real-world image dehazing. *IEEE transactions on multimedia*, 24:3934–3946, 2021. 1, 6, 7

[4] RM Drake and JE Gordon. Mie scattering. *American Journal of Physics*, 53(10):955–962, 1985. 1

[5] Earl J McCartney. Optics of the atmosphere: scattering by molecules and particles. *New York*, 1976. 1

[6] Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2341–2353, 2010. 1, 6, 7

[7] Qingsong Zhu, Jiaming Mai, and Ling Shao. A fast single image haze removal algorithm using color attenuation prior. *IEEE transactions on image processing*, 24(11):3522–3533, 2015.

[8] Raanan Fattal. Dehazing using color-lines. *ACM transactions on graphics (TOG)*, 34(1):1–14, 2014.

[9] Dana Berman, Shai Avidan, et al. Non-local image dehazing. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1674–1682, 2016.

[10] Dana Berman, Tali Treibitz, and Shai Avidan. Single image dehazing using haze-lines. *IEEE transactions on pattern analysis and machine intelligence*, 42(3):720–734, 2018. 1

[11] Wenqi Ren, Si Liu, Hua Zhang, Jinshan Pan, Xiaochun Cao, and Ming-Hsuan Yang. Single image dehazing via multi-scale convolutional neural networks. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14*, pages 154–169. Springer, 2016.

[12] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. Aod-net: All-in-one dehazing network. In *Proceedings of the IEEE international conference on computer vision*, pages 4770–4778, 2017. 2, 6, 7

[13] Dongdong Chen, Mingming He, Qingnan Fan, Jing Liao, Li-heng Zhang, Dongdong Hou, Lu Yuan, and Gang Hua. Gated context aggregation network for image dehazing and deraining. In *2019 IEEE winter conference on applications of computer vision (WACV)*, pages 1375–1383. IEEE, 2019.

[14] Haiyan Wu, Yanyun Qu, Shaohui Lin, Jian Zhou, Ruizhi Qiao, Zhizhong Zhang, Yuan Xie, and Lizhuang Ma. Contrastive learning for compact single image dehazing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10551–10560, 2021. 6, 7

[15] Tian Ye, Yunchen Zhang, Mingchao Jiang, Liang Chen, Yun Liu, Sixiang Chen, and Erkang Chen. Perceiving and modeling density for image dehazing. *European Conference on Computer Vision*, 2022. 1, 2, 3, 6, 7, 8, 10

[16] Jingang Zhang, Wenqi Ren, Shengdong Zhang, He Zhang, Yun-feng Nie, Zhe Xue, and Xiaochun Cao. Hierarchical density-aware dehazing network. *IEEE Transactions on Cybernetics*, 52(10):11187–11199, 2021. 1, 2, 6, 7

[17] Qili Deng, Ziling Huang, Chung-Chi Tsai, and Chia-Wen Lin. Hardgan: A haze-aware representation distillation gan for single image dehazing. In *European Conference on Computer Vision*, pages 722–738. Springer, 2020. 1, 3, 7

[18] Yang Yang, Chaoyue Wang, Risheng Liu, Lin Zhang, Xiaojie Guo, and Dacheng Tao. Self-augmented unpaired image de-hazing via density and depth decomposition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2037–2046, 2022. 1, 2

[19] Wenhua Lou, Yijun Li, Guowei Yang, Chenglizhao Chen, Huan Yang, and Teng Yu. Integrating haze density features for fast nighttime image dehazing. *IEEE Access*, 8:113318–113330, 2020. 1, 2

[20] Haoran Bai, Jinshan Pan, Xinguang Xiang, and Jinhui Tang. Self-guided image dehazing using progressive feature fusion. *IEEE Transactions on Image Processing*, 31:1217–1229, 2022. 2, 3, 6, 7

[21] Wei-Ting Chen, Hao-Lun Lou, Hao-Yu Fang, I-Hsiang Chen, Yi-Wen Chen, Jian-Jiun Ding, and Sy-Yen Kuo. Desmokenet: A two-stage smoke removal pipeline based on self-attentive feature consensus and multi-level contrastive regularization. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(6):3346–3359, 2021. 2, 3

[22] Xin Jin, Ling-Hao Han, Zhen Li, Chun-Le Guo, Zhi Chai, and Chongyi Li. Dnf: Decouple and feedback network for seeing in the dark. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18135–18144, 2023. 2

[23] Chun-Le Guo, Qixin Yan, Saeed Anwar, Runmin Cong, Wenqi Ren, and Chongyi Li. Image dehazing transformer with transmission-aware 3d position embedding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5812–5820, 2022. 2, 6, 7

[24] Tao Wang, Li Zhao, Pengcheng Huang, Xiaoqin Zhang, and Jiawei Xu. Haze concentration adaptive network for image dehazing. *Neurocomputing*, 439:75–85, 2021.

[25] Xin Yi, Bo Ma, Yulin Zhang, Longyao Liu, and JiaHao Wu. Two-step image dehazing with intra-domain and inter-domain adaptation. *Neurocomputing*, 485:1–11, 2022. 2

[26] Jiahao Chen, Chong Wu, Hu Chen, and Peng Cheng. Unsupervised dark-channel attention-guided cyclegan for single-image dehazing. *Sensors*, 20(21):6000, 2020. 3

[27] Ming Hong, Jianzhuang Liu, Cuihua Li, and Yanyun Qu. Uncertainty-driven dehazing network. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 906–913, 2022. 3, 6, 7

[28] Luca Bertinetto, Jack Valmadre, Joao F Henriques, Andrea Vedaldi, and Philip HS Torr. Fully-convolutional siamese networks for object tracking. In *Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part II 14*, pages 850–865. Springer, 2016. 4

[29] Anfeng He, Chong Luo, Xinmei Tian, and Wenjun Zeng. A twofold siamese network for real-time object tracking. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4834–4843, 2018.

[30] Qing Guo, Wei Feng, Ce Zhou, Rui Huang, Liang Wan, and Song Wang. Learning dynamic siamese network for visual object tracking. In *Proceedings of the IEEE international conference on computer vision*, pages 1763–1771, 2017. 4

[31] Gregory Koch, Richard Zemel, Ruslan Salakhutdinov, et al. Siamese neural networks for one-shot image recognition. In *ICML deep learning workshop*, volume 2. Lille, 2015. 4

[32] Haoran Wu, Zhiyong Xu, Jianlin Zhang, Wei Yan, and Xiao Ma. Face recognition based on convolution siamese networks. In *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pages 1–5. IEEE, 2017. 4

[33] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5728–5739, 2022. 5, 7

[34] Shibai Yin, Yibin Wang, and Yee-Hong Yang. Attentive u-recurrent encoder-decoder network for image dehazing. *Neurocomputing*, 437:143–156, 2021. 5

[35] Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao. Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 25(11):5187–5198, 2016. 6, 7

[36] Xiaohong Liu, Yongrui Ma, Zhihao Shi, and Jun Chen. Griddehazenet: Attention-based multi-scale network for image dehazing. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 7314–7323, 2019. 6, 7

[37] Hang Dong, Jinshan Pan, Lei Xiang, Zhe Hu, Xinyi Zhang, Fei Wang, and Ming-Hsuan Yang. Multi-scale boosted dehazing network with dense feature fusion. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2157–2167, 2020. 6, 7

[38] Xu Qin, Zhilin Wang, Yuanchao Bai, Xiaodong Xie, and Huizhu Jia. Ffa-net: Feature fusion attention network for single image dehazing. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 11908–11915, 2020. 6, 7

[39] Vladimir Frants, Sos Agaian, and Karen Panetta. Qcnn-h:

Single-image dehazing using quaternion neural networks. *IEEE Transactions on Cybernetics*, 2023. 6, 7

[40] Hang Sun, Bohui Li, Zhiping Dan, Wei Hu, Bo Du, Wen Yang, and Jun Wan. Multi-level feature interaction and efficient non-local information enhanced channel attention for image dehazing. *Neural Networks*, 163:10–27, 2023. 6, 7

[41] Yuda Song, Zhuqing He, Hui Qian, and Xin Du. Vision transformers for single image dehazing. *IEEE Transactions on Image Processing*, 32:1927–1941, 2023. 6, 7, 10

[42] Ye Liu, Lei Zhu, Shunda Pei, Huazhu Fu, Jing Qin, Qing Zhang, Liang Wan, and Wei Feng. From synthetic to real: Image dehazing collaborating with unlabeled real data. In *Proceedings of the 29th ACM international conference on multimedia*, pages 50–58, 2021. 6, 7, 8

[43] Codruta O Ancuti, Cosmin Ancuti, Mateu Sbert, and Radu Timofte. Dense-haze: A benchmark for image dehazing with dense-haze and haze-free images. In *2019 IEEE international conference on image processing (ICIP)*, pages 1014–1018. IEEE, 2019. 6, 7

[44] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. *CoRR*, abs/1603.08155, 2016. 6

[45] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing*, 28(1):492–505, 2018. 6

[46] Andreas Lugmayr, Martin Danelljan, and Radu Timofte. Ntire 2020 challenge on real-world image super-resolution: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 494–495, 2020. 6

[47] Codruta O. Ancuti, Cosmin Ancuti, Radu Timofte, Luc Van Gool, Lei Zhang, and Ming-Hsuan Yang. Ntire 2019 image dehazing challenge report. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2019. 6

[48] Ilya Loshchilov and Frank Hutter. Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608.03983*, 2016. 7