

Received August 17, 2021, accepted September 1, 2021, date of publication September 3, 2021, date of current version September 13, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3110428

Indirect Domain Shift for Single Image Dehazing

HUAN LIU¹, (Graduate Student Member, IEEE), AND JUN CHEN¹, (Senior Member, IEEE)

Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4L8, Canada

Corresponding author: Jun Chen (chenjun@mcmaster.ca)

ABSTRACT Despite their remarkable expressibility, convolution neural networks (CNNs) still fall short of delivering satisfactory results on single image dehazing, especially in terms of faithful recovery of fine texture details. In this paper, we argue that the inadequacy of conventional CNN-based dehazing methods can be attributed to the fact that the domain of hazy images is too far away from that of clear images, rendering it difficult to train a CNN for learning direct domain shift through an end-to-end manner and recovering texture details simultaneously. To address this issue, we propose to add explicit constraints inside a deep CNN model to guide the restoration process. In contrast to direct learning, the proposed mechanism shifts and narrows the candidate region for the estimation output via multiple confident neighborhoods. Therefore, it is capable of consolidating the expressibility of different architectures, resulting in a more accurate indirect domain shift (IDS) from the hazy images to that of clear images. We also propose two different training schemes, including hard IDS and soft IDS, which further reveal the effectiveness of the proposed method. Our extensive experimental results indicate that the dehazing method based on this mechanism dramatically outperforms the state-of-the-arts.

INDEX TERMS Single image dehazing, domain shift, deep neural network.

I. INTRODUCTION

Deep convolutional neural networks (CNNs) have been tremendously successful in many high-level computer vision tasks, e.g. image recognition [1], [2] and object detection [3], [4]. Although recent works have shown that it is also possible to learn an end-to-end CNN model for low-level vision tasks, e.g. image dehazing [5], [6], the resulting performance is still not completely satisfactory. For high-level vision tasks, it suffices to extract specific features and simply express them as very low dimensional vectors [1], which results in a relatively simple mapping. In contrast, low-level vision tasks require both global understanding of image content and local inference of texture details; as such, the associated mappings are more complicated.

One possible explanation for performance discrepancies on high-level and low-level vision tasks is as follows. For high-level vision tasks such as image recognition, a slight perturbation of the output tends to be inconsequential since the perturbed output is likely to get converted to the same one-hot vector and consequently the classification label remains unaffected. However, for low-level vision tasks such as image dehazing, any perturbation can potentially manifest in the

final result, jeopardizing the image quality. From this point of view, despite the fact that a deep CNN can in principle approximate any function, it is still difficult to train an accurate mapping that lifts the input to the target domain in one shot, since the loss function is typically very close to zero in the neighborhood of the target image [7]. We argue that a different mechanism for domain shift is needed for image dehazing, which requires both memory and understanding of image contents.

To this end, we provide explicit guidance during model optimization to lead the domain shift path across several identified confident neighborhoods, resulting in the proposed framework shown in Figure 1. More specifically, instead of only imposing the loss function on the model output, we introduce multi-scale estimation, multi-branch diversity, and adversarial loss inside the model, thereby pulling the interim outputs to specific regions then merging them in the target domain; this yields an indirect but more accurate mapping. The contributions of this paper include:

- By introducing loss functions inside a CNN model, we propose the framework of indirect domain shift (IDS) for image dehazing, which aggregates powerful expressibility of different architectures, i.e., multi-scale, multi-branch, and generator for lifting degraded images to the target domain indirectly.

The associate editor coordinating the review of this manuscript and approving it for publication was Essam A. Rashed¹.

- We provide theoretic justifications for IDS and show that it provides valuable guidance for network construction.
 - A multi-scale module takes the advantage of coarse-fine network to maintain global-local consistency.
 - A multi-branch architecture is adopted to enable precise inference of local details by providing diverse confident neighborhoods.
 - A FusionNet further improves the perceptual quality by informed ‘imagination’, rather than blindly pursuing a higher PSNR, as the multi-scale multi-branch structure has shifted degraded images close enough to the corresponding ground truth in terms of objective image quality metrics.
- It is demonstrated that IDS leads to remarkable performance improvements compared with the state-of-the-art algorithms.

II. RELATED WORKS

Image dehazing, which aims to recover a haze-free image from its hazy version, is a highly ill-posed restoration problem. The haze effect is often approximated using the atmospheric scattering model [8] given as follows:

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

where $\mathbf{I}(x)$, $\mathbf{J}(x)$, and \mathbf{A} are the observed hazy image, clear scene radiance, and global atmospheric light, respectively. The scene transmission $t(x)$ describes the portion of light that is not scattered and reaches the camera. It can be expressed as $t(x) = e^{-\beta d(x)}$, where β is the medium extinction coefficient and $d(x)$ is the depth map of pixel x .

Based on this atmospheric scattering model [8], many strategies have been proposed by taking advantage of various prior knowledge. For example, the dark channel prior [9] assumes that in non-sky patches, at least one color channel has very low intensity. The color attenuation prior [10] assumes that the image saturation decreases sharply at hazy patches, so that the difference between brightness and saturation can be utilized to estimate the haze concentration. To address the weakness of DCP for the sky region, [11] proposes to separately deal with the non-sky region and the sky region using dark channel prior and luminance stretching. In [12], the authors come up with a new color channel method to remove atmospheric scattering for single image dehazing. The overall algorithm consists of atmospheric light calculation, transmission map estimation, radiance estimation and post enhancement. Furthermore, based on the assumption that a linear relationship exists in the minimum channel between hazy and haze-free images, a fast linear-transformation-based dehazing algorithm is introduced in [13].

Recently, data-driven approaches to image dehazing have received increasing attention. [14] and [15] propose to use CNN for medium transmission estimation, which is further leveraged to recover the haze-free image. In [14], a multi-scale deep neural network is proposed to learn a mapping

between hazy images and their corresponding transmission maps. A densely connected pyramid network is proposed in [16] to jointly estimate the transmission map, atmospheric light, and dehazed images, while an effective iteration algorithm is developed in [17] to learn the haze-relevant priors. [18] further embeds the atmospheric model into the designing of CNN and proposes a feature dehazing unit to ensure end-to-end trainable. However, it is known that the atmospheric scattering model (ASM) is not valid in certain scenarios [19], which limits the applicability of the aforementioned dehazing methods. Unlike those ASM-dependent methods, [20] integrates multiple models to perform haze removal with attention, and [21] uses a GridNet-based network [22] to directly predict dehazed images via an ASM-agnostic approach. To further improve the performance in ASM-agnostic setting, [23] propose a multi-scale boosted dehazing network (MSBDN) with boosting strategy and back-projection technique. [24] firstly introduces knowledge distillation in solving dehazing problem. It allows dehazing model learn to dehaze from both ground truths and teacher outputs.

Many methods that have been developed for other image restoration tasks, e.g. deblurring, denoising, are also highly relevant. To remove blurring caused by the dynamic scenes, a multi-scale convolutional neural network is proposed in [25] to restore sharp images in an end-to-end manner. In [26], the weighted nuclear norm minimization (WNNM) problem is studied and applied to image denoising by exploiting non-local self-similarity. This work is later extended to handle arbitrary degradation, including blur and missing pixels [27]. To tackle the long-term dependency problem, the MemNet [28] is proposed by introducing a memory block, consisting of a recursive unit and a gate unit, to explicitly mine persistent memory through an adaptive learning process. To make the deep networks implementable on limited resources, a new activation unit is proposed [29], which enables the net to capture much more complex features, thus requiring a significantly smaller number of layers in order to reach the same performance. A super-resolution generative adversarial network (SRGAN) is developed in [30] to recover high-frequency details and produce more natural-looking images.

III. FORMULATION FOR INDIRECT DOMAIN SHIFT

In this section, we provide a theoretical formulation of the image dehazing problem and propose an indirect domain shift method as an effective approach to obtaining an approximation solution.

Denote the prior distribution of clear images of size $m \times n$ by p_X , which is defined on a low dimensional manifold \mathcal{M} in $\mathbb{R}^{3 \times m \times n}$. The image degradation mechanism can be modeled as a conditional distribution $p_{X|Y}$, i.e., given the clear image x , a distorted image y is generated according to $p_{Y|X}$. Note that p_X and $p_{Y|X}$ induce the joint distribution $p_{X,Y}$ as well as the conditional distribution $p_{X|Y}$; in general, both p_X and $p_{Y|X}$ need to be learned from the training data. Image dehazing can

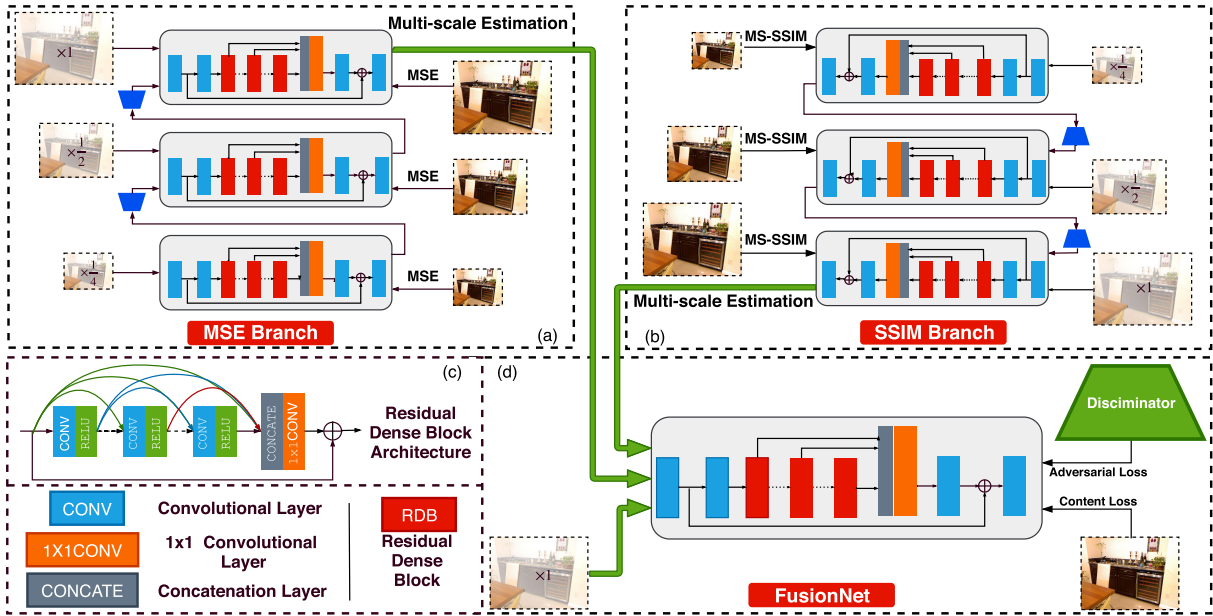


FIGURE 1. One example of the proposed IDS network. (a) and (b) are the multi-scale estimation with MSE and SSIM loss, respectively. (d) is the FusionNet with adversarial and content loss. (c) shows the legend.

be formulated as a maximum a posterior estimation problem:

$$\hat{x}_{\text{map}} = \arg \max_{\hat{x} \in \mathcal{M}} p_{X|Y}(\hat{x}|y). \quad (2)$$

In practice, one often considers the following alternative formulation:

$$\begin{aligned} \hat{x}_\ell &= \min_{\hat{x} \in \mathbb{R}^{3 \times m \times n}} \mathbb{E}[\ell(X, \hat{x})|Y = y] \\ &= \min_{\hat{x} \in \mathbb{R}^{3 \times m \times n}} \int_{\mathcal{M}} p_{X|Y}(x|y)\ell(x, \hat{x})dx, \end{aligned} \quad (3)$$

where ℓ is a loss function. In general it is expected that both \hat{x}_{map} and \hat{x}_ℓ are close to the ground truth. However, there is no guarantee that \hat{x}_ℓ belongs to \mathcal{M} .

We shall describe an IDS method, which leverages multi-scale estimation and multi-branch diversity to obtain an approximate solution of (3), then lifts it into \mathcal{M} using the adversarial loss to produce a candidate solution of (2). A network that realizes the IDS method is shown in Figure 1.

A. MULTI-SCALE ESTIMATION

Note that (3) requires the knowledge of $p_{X|Y}$, which needs to be estimated from the training data, hence we solve the following approximated version of (3), i.e.,

$$\hat{x}'_\ell = \min_{\hat{x} \in \mathbb{R}^{3 \times m \times n}} \int_{\mathcal{M}} p'_{X|Y}(x|y)\ell(x, \hat{x})dx, \quad (4)$$

where $p'_{X|Y}$ is an approximation of $p_{X|Y}$ learned from the training data. To ensure that $\hat{x}'_\ell \approx \hat{x}_\ell$ (and consequently close to the ground truth), we need $p'_{X|Y}(x|y) \approx p_{X|Y}(x|y)$ for $x \in \mathcal{M}$ (at least for x in a neighborhood of y that contains the ground truth). However, since the difference between the ground truth and the distorted version y is not negligible, this

neighborhood could be quite large, rendering a good approximation of $p_{X|Y}(\cdot|y)$ in this neighborhood difficult to obtain. Indeed, the number of parameters need to specify $p_{X|Y}(\cdot|y)$ in this neighborhood might be comparable or even larger than the available training data, hence a direct approximation can be highly unreliable, especially considering the fact that the approximation is in general done in a suboptimal way. For this reason, it is sensible to first approximate $p_{\tilde{X}|Y}$ (with \tilde{x} being a low-resolution version of the ground truth), which itself is an approximation of $p_{X|Y}$ and can be specified by a significantly smaller number of parameters (as compared to $p_{X|Y}$). In this way, we can get a good approximation of $p_{\tilde{X}|Y}$, denoted by $p'_{\tilde{X}|Y}$, and solve the following optimization problem instead:

$$\tilde{x}'_\ell = \min_{\tilde{x} \in \mathbb{R}^{3 \times m \times n}} \int_{\mathcal{M}} p'_{\tilde{X}|Y}(\tilde{x}|y)\tilde{\ell}(x, \hat{x})dx. \quad (5)$$

Since $p'_{\tilde{X}|Y}(x|y)$ is a good approximation of $p_{\tilde{X}|Y}(x|y)$, it is expected that \tilde{x}'_ℓ is close to \tilde{x} and consequently not very far away from the ground truth. Now with \tilde{x}'_ℓ at hand, we can further convert (3) to the following problem:

$$\hat{x}_\ell = \min_{\hat{x} \in \mathbb{R}^{3 \times m \times n}} \int_{\mathcal{N}(\tilde{x}'_\ell)} p_{X|\tilde{x}'_\ell, Y}(x|\tilde{x}'_\ell, y)\ell(x, \hat{x})dx, \quad (6)$$

where $\mathcal{N}(\tilde{x}'_\ell)$ is a neighborhood of \tilde{x}'_ℓ that is large enough to cover the ground truth. It suffices to have a good approximation $p_{X|\tilde{x}'_\ell, Y}(\cdot|\tilde{x}'_\ell, y)$ over $\mathcal{N}(\tilde{x}'_\ell)$. The above procedure is repeated until the required neighborhood is small enough.

We assume that the smaller the neighborhood becomes, the fewer number of parameters are needed to specify the distribution defined over this neighborhood and consequently the approximation becomes easier. Multi-scale estimation is introduced to mimic conventional coarse-to-fine optimization

methods and has been widely applied in many computer vision tasks [14], [25], [31], [32].

B. MULTI-BRANCH DIVERSITY

The idea underlying multi-branch diversity is similar. Suppose we adopt two branches with different loss functions, denoted by ℓ_1 and ℓ_2 , respectively, then (6) becomes

$$\hat{x}_\ell = \min_{\hat{x} \in \mathbb{R}^{3 \times m \times n}} \int_{\mathcal{N}(\tilde{x}_{\ell_1}) \cap \mathcal{N}(\tilde{x}_{\ell_2})} \times P_{X|\tilde{X}_{\ell_1}, \tilde{X}_{\ell_2}, Y}(x|\tilde{x}_{\ell_1}, \tilde{x}_{\ell_2}, y)\ell(x, \hat{x})dx. \quad (7)$$

It should be clear that multi-branch diversity further narrows the region over which the distribution needs to be estimated. In our experiments, we choose ℓ_1 and ℓ_2 to be mean square error (MSE) and structural similarity index (SSIM) loss, respectively. The reason we choose MSE and SSIM as loss functions is that MSE focuses on the pixel-level difference while SSIM pays more attention to the perceptual quality. See Figure 1 (a) and (b) for the architecture of two multi-scale estimation branches of the proposed IDS network.

C. ADVERSARIAL LOSS

The role of the adversarial loss ℓ_{ad} is to lift \hat{x}_ℓ into \mathcal{M} . Specifically, consider a neural network subject to the weighted loss $\ell + \lambda\ell_{ad}$, which can be interpreted as solve the following problem:

$$\hat{x}_{\ell+\lambda\ell_{ad}} = \arg \max_{\hat{x} \in \mathcal{N}(\hat{x}_\ell, \lambda)} p_X(\hat{x}), \quad (8)$$

where $\mathcal{N}(\hat{x}_\ell, \lambda)$ is a neighborhood of \hat{x}_ℓ . In general, this optimization problem tends to give a reconstruction that falls into \mathcal{M} since p_X is only positive on \mathcal{M} . Note that the size of $\mathcal{N}(\hat{x}_\ell, \lambda)$ depends on λ . Specifically, $\mathcal{N}(\hat{x}_\ell, \lambda)$ is large when λ is large. In the extreme case of $\lambda \rightarrow \infty$, we have $\hat{x}_{\ell+\lambda\ell_{ad}} \rightarrow \arg \max_{\hat{x} \in \mathcal{M}} p_X(\hat{x})$; while when λ is very small, $\mathcal{N}(\hat{x}_\ell, \lambda)$ may have no intersection with \mathcal{M} , and in this case (8) reduces to (3). In principle it is desirable to choose the smallest λ such that $\mathcal{N}(\hat{x}_\ell, \lambda)$ intersects with \mathcal{M} . It is also worth noting that p_X is in general unknown. So one has to solve a modified version of (8) with p_X replaced by p'_X , which is an approximation of p_X learned from the training data.

The adversarial loss serves an important role of generating texture details in image restoration. One of the reasons for its success in our framework is that, by leveraging multi-scale estimation and multi-branch diversity, one can already obtain an good estimate \hat{x}_ℓ which is in a narrow neighboring region of \mathcal{M} , and consequently the generator does not need much “imagination” to produce a natural-looking image. However, we observe the similar phenomenon reported in [30] that adversarial loss is helpful for faithful reproduction, even though the final PSNR metric is slightly lower. Nevertheless, we introduce the adversarial loss to obtain better perceptual quality but not expect higher PSNR value. The relevant ablation study can be found in Section V-C.

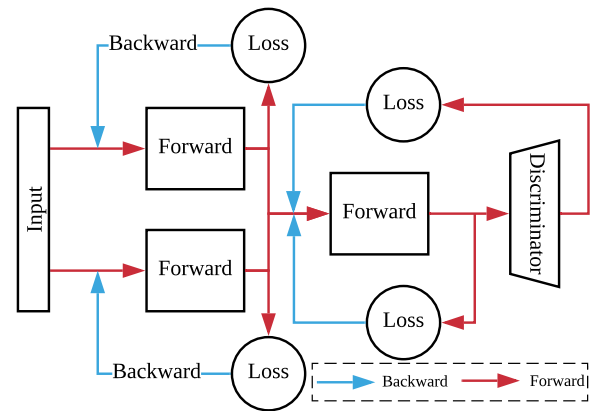


FIGURE 2. The isolated training of one iteration in hard IDS.

IV. IMPLEMENTATION

In this section, we provide a detailed implementation of the indirect domain shift (IDS). We also propose two training schemes, i.e., the hard IDS and soft IDS.

A. NETWORK ARCHITECTURE

The proposed IDS network is shown in Figure 1, which consists of three basic components, i.e., the MSE branch, the MS-SSIM branch, and the FusionNet. The MSE and SSIM branches are built with multi-scale structure to successively map hazy images to their clear counterpart at different resolution levels (as in (6)); moreover, they are supervised by non-identical loss functions to ensure differentiated outputs. The FusionNet completes the domain shift process by merging the outputs from the two branches together with the input hazy image into a single clear image (as in (7)). We train the FusionNet (see Figure 1 (d)) using a content loss defined as the weighted sum of MSE loss and perceptual loss [33]. The weight is carefully selected by searching from 1.0, 10^{-1} , 10^{-2} , and 10^{-3} . We find that our network achieves the best performance when the weight is set to 10^{-2} . An adversarial loss (see (8)) is also imposed on the FusionNet to enhance the perceptual quality of the final result.

To be specific, inside each diversity branch, there are three sub-networks, each performing domain shift at a different scale level. The input of the coarse-scale sub-network is obtained from the original hazy image via bi-linear interpolation with a down-sampling factor of 4. Its output is up-sampled with a factor of 2 via pixel shuffle [34], then fed into the medium-scale sub-network, together with the down-sampled hazy image representation by a factor of 2. The input of the fine-scale sub-network is the concatenation of the original hazy image representation and the up-sampled output of medium-scale sub-network.

It is known that residual networks (ResNets) can facilitate gradient flow while dense networks (DenseNets) help maximize the use of feature layers via concatenation and dense connection. To capitalize on their respective strengths, [35] proposes so-called residual dense networks (RDNs), which

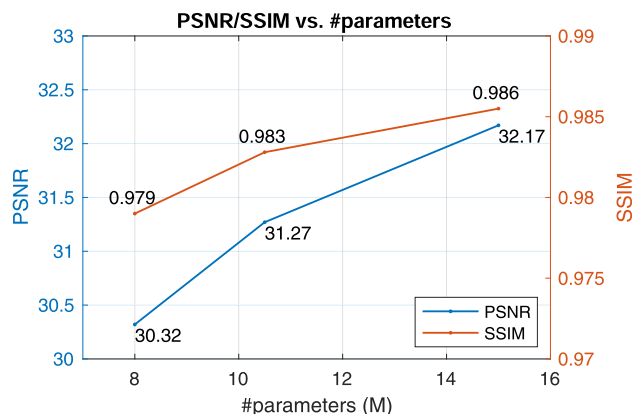


FIGURE 3. The performance of hard IDS with different parameters.

consist of contiguous memory blocks, local residual learning blocks and global feature fusion blocks.

In this work, we use RDNs as the fundamental building components of the proposed IDS network. See Table 1 for detailed specifications. Note that hard IDS and soft IDS adopt the same network structure, but differ in terms of the number of trainable parameters. Model depth will be detailed in Section V-D.

B. TRAINING SCHEME

To handle the coexistence of multiple loss functions, we propose two back-propagation strategies characterized by different effective ranges of the loss functions. Specifically, we can separately update each module according to the associated loss function or jointly update all modules according to a global loss that aggregates the local ones. This results in the two IDS training schemes, i.e., hard IDS and soft IDS.

1) HARD IDS

We first present the isolated training strategy for hard IDS shown in Figure 2. Specifically, each module is supervised independently by the associated loss functions and deliver dehazed images to the next stage after updating their weights. Note that in this case, the convergence of the entire network does not depend on the convergence of all loss functions, which means that the network performance may become stable before all loss functions are small enough. This is a consequence of direct mapping, since for each mapping step it suffices to enter one of many (almost) equally good confident neighborhoods, resulting in lower computational load. One advantage of isolated updating is that the gradient vanishing problem can be alleviated. Recall that this problem is caused by the emergence of small gradients in the earlier layers of very deep networks during back-propagation. As a comparison, isolated training shortens the back-propagation path, but maintains the depth of forward inference, at the expense of heterogeneous convergence rates of different loss functions. It is also worth noting that the isolated training strategy closely follows our analytical formulation which dictates how to shift from one domain to another. Therefore,

TABLE 1. The configuration of the shadow, medium, and deep Hard IDS corresponding to Figure 3.

#RDB/#Conv		Shadow	Medium	Deep
Branch	Coarse	4/3	5/3	6/3
	Finer	6/4	7/4	8/4
	Finest	8/5	9/5	10/5
FusionNet		10/6	12/6	15/6

the success of hard IDS can be viewed as a good indication of the correctness of our theoretical framework.

2) SOFT IDS

In contrast to hard IDS, here a global loss function obtained by combining all local module losses is used to update network parameters via end-to-end back-propagation. Although the local losses are evaluated based on the images output by the respective modules, only the feature map from the penultimate convolutional layer of each module is delivered to the next module. This enables soft IDS to accomplish the desired task largely in the feature space. The fact that each module no longer has to re-map the previous module's output images back to the feature space is helpful for reducing the number of parameters and also making the indirect shifting path 'smoother'. Another advantage of soft IDS is that there is no need to be concerned with the convergence of a specific module as in hard IDS, which facilitates the training process.

In summary, the differences between Hard and Soft IDS are in two main aspects: (1) As in Figure 4, Hard IDS and Soft IDS deliver images and features to the next stages, respectively. (2) The Hard IDS adopts isolated training (optimization over modules independently), while Soft IDS computes the summation of all the local module losses and optimizes the entire network in an iteration.

V. ABLATION STUDY

We conduct ablation studies to investigate the respective contributions of multi-scale estimation, multi-branch diversity, and adversarial loss using RESIDE-standard indoor dataset [36] that will be introduced in detail in Section VI-A. To eliminate the influence of other factors, all training configurations are kept the same as that presented in Section VI-B, including the total number of trainable parameters for each network. More detailed analysis is shown in supplementary.

A. MULTI-SCALE ESTIMATION

As mentioned in Section III-A, a direct mapping can be highly unreliable, since the number of trainable parameters might be comparable or even larger than the available training data. To overcome this problem, a multi-scale network is applied in the first stage of IDS. Another important property of such coarse-to-fine estimation is the local-global consistency: the coarse-scale network first estimates the holistic structure of the image scene, and then a fine-scale network

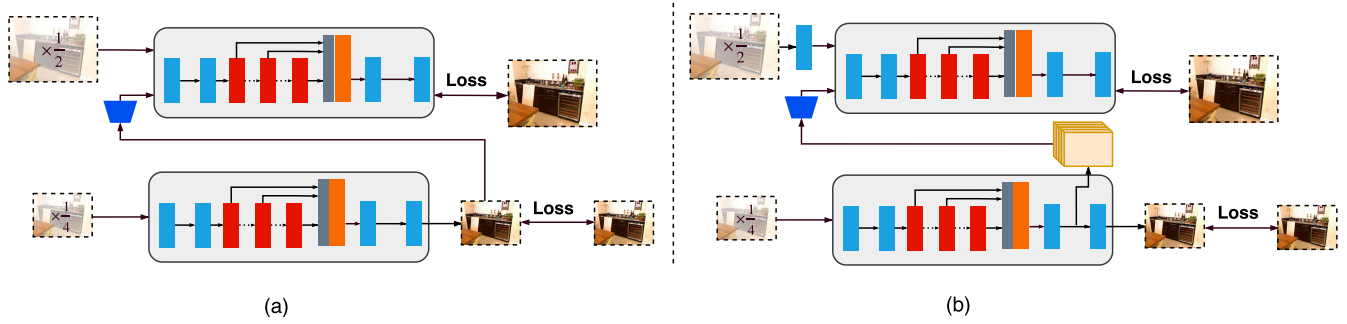


FIGURE 4. The difference between (a) Hard IDS and (b) Soft IDS.

TABLE 2. Ablation studies on the SSIM/PSNR performance. The best performance is shown in bold, while second best results are with underline.

(a) Hard IDS					(b) Soft IDS				
Scale	Branch	Adversarial	PSNR	SSIM	Scale	Branch	Adversarial	PSNR	SSIM
X	✓	✓	31.13	0.983	X	✓	✓	34.12	0.985
✓	X	✓	29.80	0.982	✓	X	✓	32.75	0.984
✓	✓	X	31.32	0.986	✓	✓	X	34.74	<u>0.986</u>
X	X	✓	30.55	0.975	X	X	✓	33.92	0.981
✓	✓	✓	<u>32.17</u>	0.986	✓	✓	✓	34.74	0.987

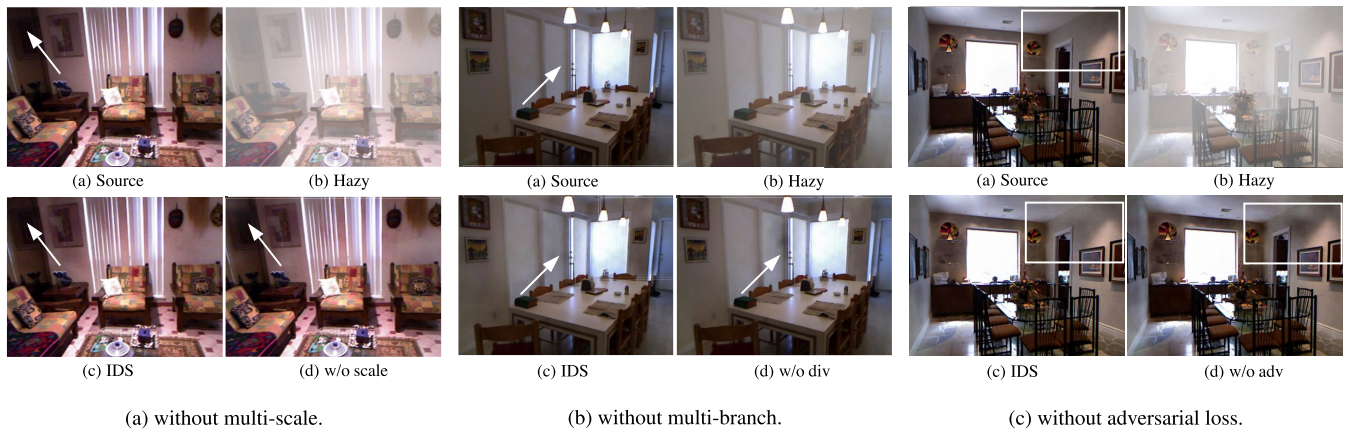


FIGURE 5. Some output examples of Hard IDS without multi-scale estimation (w/o scale), without multi-branch diversity (w/o div), only with adversarial loss (o/w adv), and without adversarial loss (w/o adv) in the ablation study, respectively.

performs refinement based on both local information and the coarse global estimation. To further study the influence of such coarse-to-fine structure, we test the performance of IDS framework without multi-scale estimation (w/o scale).

Following the ablation principle, we remove the coarse-scale network and make the fine-scale network deeper to have the same number of parameters. One output example is presented in Figure 5a indicating that hard IDS w/o scale is able to recover the image reasonably well, but with some local inconsistency: the haze at the up-left corner is not removed faithfully. This verifies the above analysis that multi-scale network is able to capture both local and global features. We present the performance on PSNR and SSIM for both hard IDS and soft IDS in Table 2 (a) and (b), respectively. It can be seen that IDS w/o scale performs worse than IDS (especially

in soft IDS), indicating that the local inconsistency has impact on both the quantitative metrics and perceptual quality.

B. MULTI-BRANCH DIVERSITY

Using multi-scale estimation with MSE loss, one can realize domain shift to a certain extent. However, some important information may get lost along the way. To keep the information diversity, we introduce one more multi-scale branch and employ SSIM loss in this branch. This strategy enables a more precise inference of local details by providing distinctive confident neighborhoods identified by different branches. To further illustrate its effectiveness of this strategy, we test the performance of IDS without multi-branch diversity (w/o div).

Similarly, we remove the second branch and make the first branch deeper. One of the examples is presented in Figure 5b, in which the IDS w/o div sometimes delivers erroneous detail inference, since the “dark area” between the light and the wall clearly should not exist. This is further verified by the overall validation shown in the Table 2, in which there is a large performance gap between IDS and IDS w/o div, indicating that it is well worth having two branches.

C. ADVERSARIAL LOSS

The adversarial loss (together with the content loss) is employed at the last stage (i.e., the FusionNet) of the proposed IDS framework and is served to obtain high visual quality. The FusionNet takes the estimates from the two branches, in conjunction with the original hazy image, as the input and generates the final output with perceptually satisfactory high-frequency details via proper fusion. Since the estimates produced by the two branches are already in the neighboring domains of the target, the generator does not need to rely on pure “imagination” to create texture details; instead, it could, to a great extent, maintain the perceptual reality rather than blindly pursue a higher PSNR [30].

To prove this, we show that IDS without adversarial loss is able to produce a high PSNR but NOT able to obtain better perceptual quality. Following the ablation principle, we construct IDS without adversarial loss (w/o adv) by simply removing discriminator. As can be seen, IDS w/o adv produces a slightly higher PSNR in Figure 5c (26.508), but obviously lower perceptual quality than IDS (26.094), as the wall is printed “darker” partially to minimize the MSE distance. This demonstrates the generalization capability of the generator and provides further justifications for the IDS framework.

To further prove the necessity of adversarial loss, we compare with GridDehaze [21]. GridDehaze [21] is a pure CNN based dehazing method without adopting adversarial loss to generate natural distributed outputs. From Figure 6, it shows that the generated images from Soft IDS tend to be closer to the ground truth with less inconsistent color gradients on the road, sky, and wall. This verifies the phenomenon that the adversarial loss is introduced to obtain better perceptual quality but not blindly pursue higher PSNR value.

D. MODEL DEPTH

This section is devoted to investigating the impact of model depth on the performance of our hard IDS method. By adjusting the number of convolutional and dense residual blocks, we construct shadow, medium, and deep models with 8 M, 10.5 M, and 15 M trainable parameters, respectively. Detailed specifications is shown in Table 1. As expected, the deep model achieves the best overall performance in terms of both PSNR and SSIM. As illustrated in Figure 3, both PSNR and SSIM values improve dramatically as the number of parameters increases, which further verifies the effectiveness of the IDS framework. It is worth mentioning that albeit with fewer

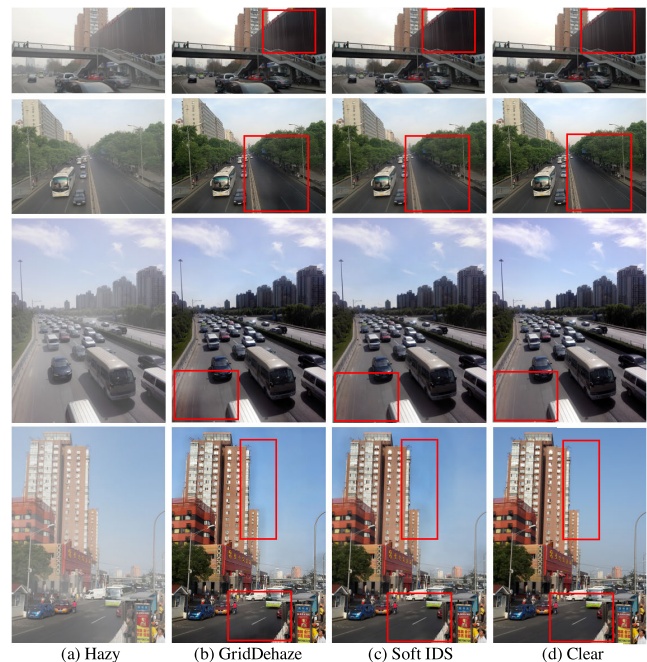


FIGURE 6. The output examples from SOTS outdoor testing set.

trainable parameters (around 4.3 M), soft IDS still manages to outperform hard IDS as shown in Table 3.

VI. EXPERIMENTS

In this section, we further compare the proposed IDS network with several state-of-the-art dehazing algorithms, including dark channel prior (DCP) [9], DehazeNet [15], AOD-Net [37], gated fusion network (GFN) [38], GridDehazeNet (GridDehaze) [21], PFD [18] and MSBDN [23]. For a fair comparison, all these algorithms are evaluated on both synthetic and realistic datasets in terms of visual effect and quantitative accuracy. We adopt the peak signal to noise ratio (PSNR) [39] and the structural similarity index (SSIM) [40] for evaluation.

A. BENCHMARK DATASET

For training and testing purposes, we use the RESIDE-standard dataset [36], which is a benchmark for single image dehazing. The indoor training set (ITS) of RESIDE-standard contains 13990 synthetic hazy indoor images (together with haze-free counterparts). These synthetic images are generated using NYU2 [41] and Middlebury stereo [42] with the medium extinction coefficient β chosen uniformly from (0.6, 1.8) and the global atmospheric light \mathbf{A} chosen uniformly from (0.7, 1.0). The outdoor training set (OTS) of RESIDE-standard contains 296695 hazy images generated from 8477 clear counterparts with β chosen uniformly from (0.04, 0.2) and \mathbf{A} chosen uniformly from (0.8, 1.0). The testing set (SOTS) of RESIDE-standard contains 500 synthetic hazy indoor/outdoor images (together with haze-free counterparts). We also perform comparisons using the real-world hazy image dataset in [43] to show the perceptual difference.

TABLE 3. The SSIM/PSNR performance of different methods on SOTS-indoor, and SOTS-outdoor. Our proposed methods and improved network with RCAN outperform the others.

Dataset	Metrics	DCP	DehazeNet	AOD-Net	GFN	GridDehaze	PFD	MSBDN	Hard IDS	Soft IDS	RCAN IDS
Indoor	PSNR	16.62	21.14	19.06	22.30	32.16	32.68	33.79	32.17	34.74	35.34
	SSIM	0.8179	0.8472	0.8504	0.880	0.9836	0.9760	0.9840	0.9860	0.9869	0.9901
Outdoor	PSNR	19.13	24.75	24.14	28.29	30.86	31.17	31.33	30.78	31.52	32.73
	SSIM	0.8605	0.9269	0.9198	0.9621	0.9819	0.9825	0.9832	0.9815	0.9832	0.9873

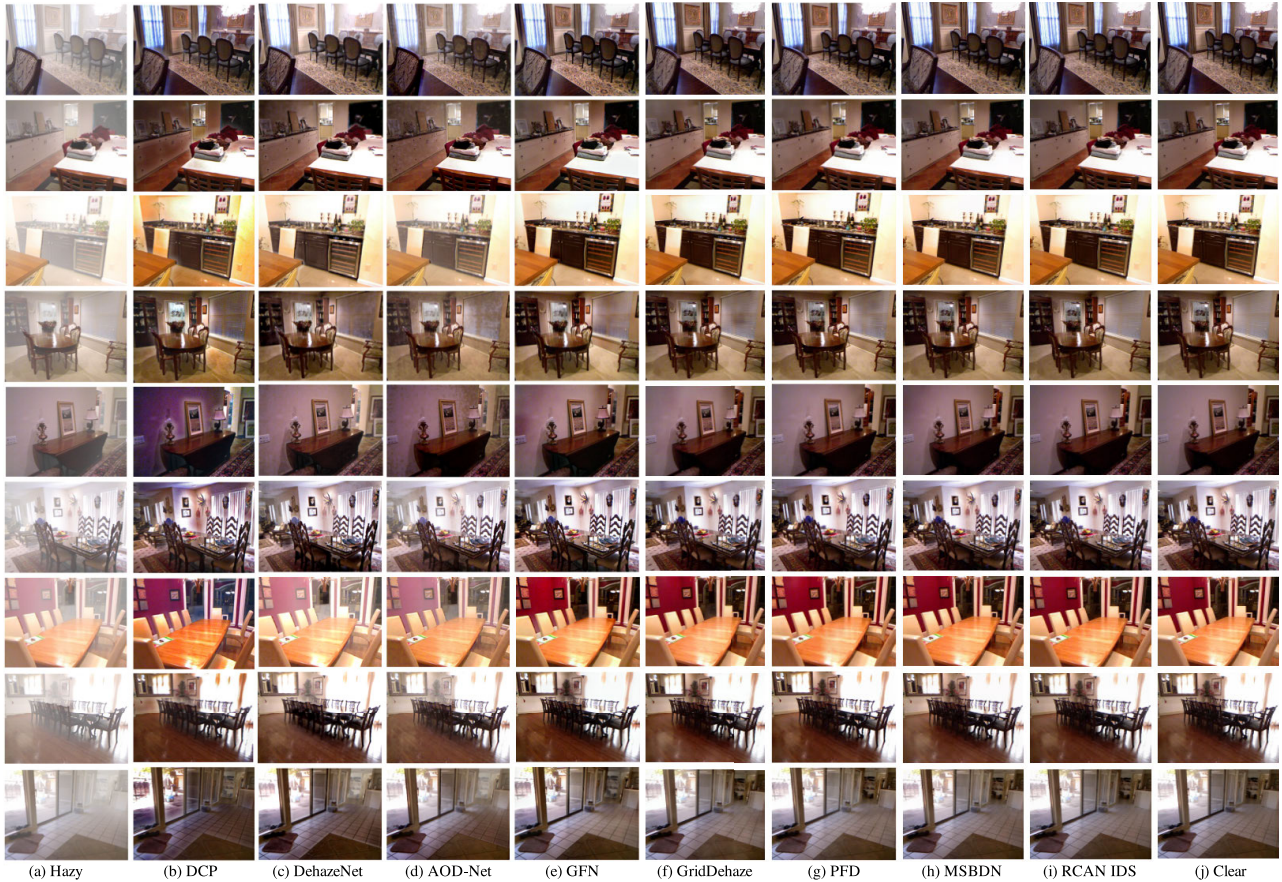


FIGURE 7. The output examples from SOTS indoor testing set of the SOTA methods.

B. TRAINING DETAILS

Our algorithm is implemented using the PyTorch library [44] and all tests are conducted on the same GPU of Nvidia Titan Xp. We train the network with the following configuration: the Adam optimizer [45] is applied with $\beta_1 = 0.9$ and $\beta_2 = 0.999$, where a mini-batch size of 10, a patch size of 180×180 , an initial learning rate of 10^{-4} are adopted. For hard IDS, the learning rate decays with a multiplicative factor of 0.5 every 120 epochs for a total of 700 epochs, while soft IDS is trained for 100 epochs with the learning rate reduced by half on the 60th, the 80th, and the 90th epochs. Besides, horizontal/vertical random flipping is applied for data augmentation. It is worth mentioning that after random flipping of both input and target images, the training data are still paired. Therefore, such an augmentation strategy is not

harmful to supervised training but help expand the size of training data.

C. RCAN AS SUBSTITUTE

The proposed IDS framework is generic in nature and admits many different concrete implementations. In this work, we have focused on a particular implementation with RDNs as fundamental building blocks. However, this is by no means the best possible one. Indeed, the performance of our IDS network can be further improved by adopting more powerful substitutes of RDNs. To demonstrate this, we replace RDNs in soft IDS by residual channel attention networks (RCANs) [46] with the same number of trainable parameters. We further illustrate the effectiveness of adopting RCANs as substitute in the following experimental results.

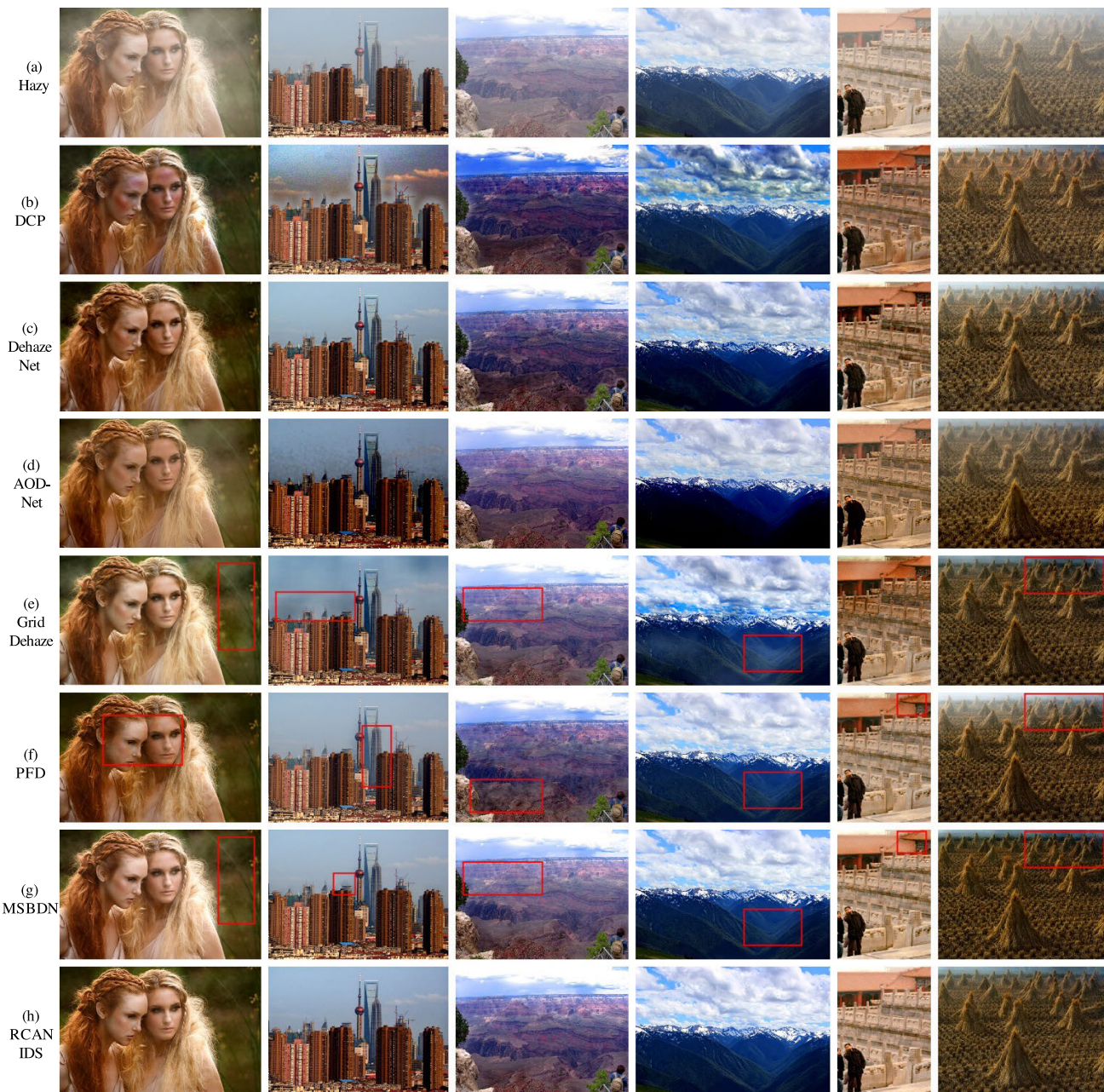


FIGURE 8. The output examples from real-world images in Fattal et.al. [43] to compare with SOTA DNN based methods.

D. EVALUATION ON BENCHMARK DATASET

We train our network from scratch on RESIDE-standard ITS, OTS and validate it on the separated testing dataset SOTS. The quantitative results and the qualitative results are shown in Table 3 and Figure 7, respectively. Here hard IDS corresponds to the deep model in Table 1, while soft IDS is as described in Section V-D. It can be seen from Table 3 that soft IDS outperforms the other methods under comparison in terms of PSNR and SSIM. In particular, the PSNR achieved by soft IDS reaches 34.74 on SOTS indoor dataset. Moreover, with the boost from RCANs substitute, RCAN IDS outperforms the others by a large margin.

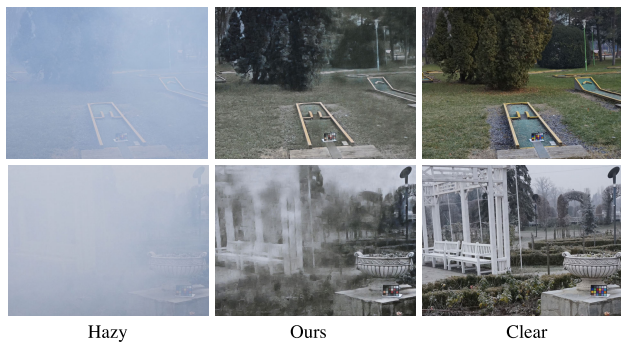
As for visual quality, prior-based methods [9] overestimate the haze thickness, which results in color distortion (e.g. the color of the wall turns purple in the fifth row in Figure 7). Although some learning-based baseline methods [15], [37] avoid the color distortion problem, they tend to deliver unsatisfactory haze removal results for shaded regions. For example, in the seventh row of Figure 7, the area behind the arch should be dark; however, the restoration results produced by most baseline methods show light color instead. This is probably because of that the baseline methods fail to correctly estimate the depth information and consequently mislead by the haze effect. GFN generates decent results, and removes

TABLE 4. Average per-image (620 × 460) runtime (second) on SOTS-indoor.

DNN Based (GPU Running Time)								
DehazeNet	AOD-Net	GFN	GridDehaze	PFD	MSBDN	Hard IDS	Soft IDS	RCAN IDS
0.190s	0.004s	0.011s	0.22s	0.103	0.088	0.048s	0.035s	0.041s



(a) The output samples from O-Haze testing set.



(b) The output samples from Dense-Haze testing set.

FIGURE 9. Qualitative evaluation on Dense-Haze [47] and NH-Haze [48] dataset.

the haze in this area reasonably well. A possible explanation is that GFN does not rely on depth estimation for haze removal; it can also be attributed to the multi-scale approach adopted by GFN, which is an important ingredient of the IDS framework as well. Exploiting the full strength of IDS enables us to obtain better dehazing results. GridDehaze [21], PFD [18] and MSBDN [23] are the methods that can produce dehazed images comparable to ours. However, they still generate inconsistent color gradients on the venetian blinds in the fourth row. On the other hand, it can be seen in Figure 7 that our dehazed images can hardly be distinguished from the ground truth.

E. EVALUATION ON REAL-WORLD PHOTOGRAPHS

We further show the dehazing results on real-world images in [43] to illustrate the generalization ability of IDS. In Figure 8, Prior-based method [9] introduces color distortion and over enhancement on images.

TABLE 5. The SSIM/PSNR performance of different methods on O-Haze and Dense-Haze dataset. Our proposed methods outperform the others.

Metrics	O-HAZE		Dense-Haze		
	PSNR	SSIM	Metrics	PSNR	SSIM
DCP	12.92	0.505	DCP	10.85	0.404
AOD-Net	17.69	0.616	AOD-Net	13.30	0.469
GridDehaze	22.76	0.721	GridDehaze	14.56	0.493
MSBDN	23.28	0.743	MSBDN	15.18	0.509
Ours	23.84	0.766	Ours	15.78	0.543

It is clear that DehazeNet [15], and AOD-Net [37] fail to remove haze completely, especially in the last column where heavy haze can still be seen around the haystack. Moreover, they also tend to over-enhance the images (e.g. the mountains in the fourth column). Although GridDehaze [21], PFD [18] and MSBDN [23] work well on the synthetic dataset, its generalization performance on real images is unsatisfactory. The red boxes in Figure 8 locate their unsatisfactory regions. Their weaknesses include color distortion, incomplete haze removal and over enhancement. We also notice that the proposed IDS is able to not only remove haze successfully, regardless whether it is dense or light, but also restore the texture details faithfully, which further proves the effectiveness of our method.

F. EVALUATION ON REAL-WORLD DATASETS

The evaluation is conducted on the O-Haze [49], and Dense-Haze [47] datasets. The Two real-world datasets is challenging since they contain limited training images (45 and 55 respectively) and vivid haze patterns. Therefore, the performance on the two dataset can be a good indication to the effectiveness of the proposed methods. The training on the two datasets adopts same strategies as introduced in Section VI-B. For fair comparison, we omit to use pre-trained weights or data augmentations that are not introduced in Section VI-B. We demonstrate the evaluation quantitatively and qualitatively in Table 5 and Figure 9.

1) RESULTS ON NTIRE2018 O-HAZE

We evaluate our proposed IDS on O-Haze dataset [49] following the data split in official NTIRE2018-Dehazing challenge [50]. It can be observed in Table 5 that our IDS outperforms the other methods in terms of PSNR and SSIM. Figure 9a shows that our approach reconstructs faithful and sharp haze free images with good perceptual quality.

2) RESULTS ON NTIRE2019 DENSE-HAZE

In contrast to O-Haze that mostly contains light haze, Dense-Haze [47] records images with denser and more

homogeneous haze layer. We follow NTIRE-2019 challenge [51] to conduct evaluation. Qualitative results in Figure 9b demonstrate that even if the background scene is occluded by thick haze, our IDS is still able to restore these regions. In particular, since the second testing sample in Figure 9b is covered by severe haze, the background scene is almost invisible to human eyes. Nevertheless, our IDS surprisingly removes dense haze and reconstructs identifiable details. Quantitative comparisons in Table 5 illustrate that our IDS is the top performing method.

G. RUNTIME

Table 4 shows runtime comparisons on the SOTS dataset. Our method is ranked the third among DNN-based methods. It is worth mentioning that in our implementation multi-scale estimation is performed branch by branch. A significant reduction in runtime is possible via a parallel implementation of multi-scale estimation in two branches.

VII. CONCLUSION

In this paper, it is shown that the traditional direct mapping methods cannot provide accurate direct mapping for image dehazing. To solve this problem, an indirect domain shift (IDS) method is proposed by adding explicit loss functions inside a deep CNN model to guide the dehazing process. Multi-scale estimation, multi-branch diversity, and adversarial loss play important roles in this method as shown by the ablation studies. We also propose two training schemes, which have their respective advantages. Specifically, hard IDS is less demanding in terms of computational resources and alleviates the gradient vanishing problem. Besides, hard IDS is designed according to our theoretical formulation and its success provides a strong empirical indication of the correctness of our indirect domain shift mechanism. On the other hand, soft IDS is easier to implement and in general yields better performance. We show that IDS achieves remarkable improvements compared with the state-of-the-art on five dehazing datasets. Despite the success of our method, the visual performance of IDS is not completely satisfactory on Dense-Haze dataset. Since the deep learning methods often require large-scale datasets for training, we believe the performance of our method on Dense-Haze dataset can be further improved by simply acquiring more training samples. From another perspective, one interesting direction for our future work is to enhance the IDS framework to enable good generalization with limited training data.

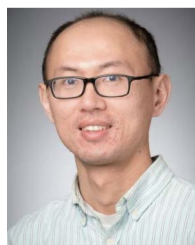
REFERENCES

- [1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Conf. Neural Inf. Process. Syst.*, 2012, pp. 1–9.
- [2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [3] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1440–1448.
- [4] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.
- [5] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "DehazeNet: An end-to-end system for single image haze removal," *IEEE Trans. Image Process.*, vol. 25, no. 11, pp. 5187–5198, Nov. 2016.
- [6] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. Eur. Conf. Comput. Vis. Cham, Switzerland: Springer*, 2016, pp. 694–711.
- [7] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2999–3007.
- [8] S. G. Narasimhan and S. K. Nayar, "Vision and the atmosphere," *Int. J. Comput. Vis.*, vol. 48, no. 3, pp. 233–254, 2002.
- [9] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341–2353, Dec. 2011.
- [10] Q. Zhu, J. Mai, and L. Shao, "A fast single image haze removal algorithm using color attenuation prior," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3522–3533, Nov. 2015.
- [11] G. Sahu and A. Seal, "Image dehazing based on luminance stretching," in *Proc. Int. Conf. Technol. (ICIT)*, Dec. 2019, pp. 388–393.
- [12] G. Sahu, A. Seal, O. Krejcar, and A. Yazidi, "Single image dehazing using a new color channel," *J. Vis. Commun. Image Represent.*, vol. 74, Jan. 2021, Art. no. 103008.
- [13] W. Wang, X. Yuan, X. Wu, and Y. Liu, "Fast image dehazing method based on linear transformation," *IEEE Trans. Multimedia*, vol. 19, no. 6, pp. 1142–1155, Jun. 2017.
- [14] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, "Single image dehazing via multi-scale convolutional neural networks," in *Proc. Eur. Conf. Comput. Vis. Springer*, 2016, pp. 154–169.
- [15] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "DehazeNet: An end-to-end system for single image haze removal," *IEEE Trans. Image Process.*, vol. 25, no. 11, pp. 5187–5198, Nov. 2016.
- [16] H. Zhang and V. M. Patel, "Densely connected pyramid dehazing network," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3194–3203.
- [17] Y. Liu, J. Pan, J. Ren, and Z. Su, "Learning deep priors for image dehazing," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 2492–2500.
- [18] J. Dong and J. Pan, "Physics-based feature dehazing networks," in *Proc. Eur. Conf. Comput. Vis. Springer*, 2020, pp. 188–204.
- [19] Y. Li, R. T. Tan, and M. S. Brown, "Nighttime haze removal with glow and multiple light colors," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 226–234.
- [20] Z. Deng, L. Zhu, X. Hu, C.-W. Fu, X. Xu, Q. Zhang, J. Qin, and P.-A. Heng, "Deep multi-model fusion for single-image dehazing," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 2453–2462.
- [21] X. Liu, Y. Ma, Z. Shi, and J. Chen, "GridDehazeNet: Attention-based multi-scale network for image dehazing," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 7314–7323.
- [22] D. Fourure, R. Emonet, E. Fromont, D. Muselet, A. Tremeau, and C. Wolf, "Residual conv-deconv grid network for semantic segmentation," 2017, *arXiv:1707.07958*. [Online]. Available: <http://arxiv.org/abs/1707.07958>
- [23] H. Dong, J. Pan, L. Xiang, Z. Hu, X. Zhang, F. Wang, and M.-H. Yang, "Multi-scale boosted dehazing network with dense feature fusion," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 2157–2167.
- [24] M. Hong, Y. Xie, C. Li, and Y. Qu, "Distilling image dehazing with heterogeneous task imitation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 3462–3471.
- [25] S. Nah, T. H. Kim, and K. M. Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 257–265.
- [26] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2862–2869.
- [27] N. Yair and T. Michaeli, "Multi-scale weighted nuclear norm image restoration," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3165–3174.
- [28] Y. Tai, J. Yang, X. Liu, and C. Xu, "MemNet: A persistent memory network for image restoration," in *Proc. Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 4539–4547.
- [29] I. Kligvasser, T. R. Shaham, and T. Michaeli, "xUnit: Learning a spatial activation function for efficient image restoration," in *Proc. CVPR*, Jun. 2018, pp. 2433–2442.

- [30] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 105–114.
- [31] D. Eigen, C. Puhrsch, and R. Fergus, "Depth map prediction from a single image using a multi-scale deep network," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2366–2374.
- [32] D. Eigen and R. Fergus, "Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 2650–2658.
- [33] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. ECCV*, 2016, pp. 694–711.
- [34] W. Shi, J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1874–1883.
- [35] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 2472–2481.
- [36] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang, "Benchmarking single-image dehazing and beyond," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 492–505, Jan. 2019.
- [37] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "AOD-net: All-in-one dehazing network," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4770–4778.
- [38] W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W. Liu, and M.-H. Yang, "Gated fusion network for single image dehazing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3253–3261.
- [39] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss functions for image restoration with neural networks," *IEEE Trans. Comput. Imag.*, vol. 3, no. 1, pp. 47–57, Mar. 2017.
- [40] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [41] P. K. N. Silberman, D. Hoiem, and R. Fergus, "Indoor segmentation and support inference from RGBD images," in *Proc. ECCV*, 2012, pp. 746–760.
- [42] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, Jun. 2003, pp. I–I.
- [43] R. Fattal, "Dehazing using color-lines," *ACM Trans. Graph.*, vol. 34, no. 1, p. 13, Dec. 2014.
- [44] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," in *Proc. NIPS Workshop*, 2017.
- [45] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [46] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 286–301.
- [47] C. O. Ancuti, C. Ancuti, M. Sbert, and R. Timofte, "Dense-haze: A benchmark for image dehazing with dense-haze and haze-free images," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 1014–1018.
- [48] C. O. Ancuti, C. Ancuti, and R. Timofte, "NH-HAZE: An image dehazing benchmark with non-homogeneous hazy and haze-free images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 444–445.
- [49] C. O. Ancuti, C. Ancuti, R. Timofte, and C. De Vleeschouwer, "O-HAZE: A dehazing benchmark with real hazy and haze-free outdoor images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 754–762.
- [50] C. Ancuti, C. O. Ancuti, and R. Timofte, "NTIRE 2018 challenge on image dehazing: Methods and results," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) Workshops*, Jun. 2018, pp. 891–901.
- [51] C. O. Ancuti, C. Ancuti, R. Timofte, L. V. Gool, L. Zhang, and M.-H. Yang, "NTIRE 2019 image dehazing challenge report," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2019, pp. 2241–2253.



HUAN LIU (Graduate Student Member, IEEE) received the B.Eng. degree in communication engineering from the University of Electronic Science and Technology of China (UESTC), in 2017. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, McMaster University. His research interests include computer vision and machine learning.



JUN CHEN (Senior Member, IEEE) received the B.S. degree in electronic engineering from Shanghai Jiao Tong University, in 2001, and the M.S. and Ph.D. degrees in electrical and computer engineering from Cornell University, in 2004 and 2006, respectively.

He was a Postdoctoral Research Associate with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, from September 2005 to July 2006, and a Postdoctoral Fellow with IBM Thomas J. Watson Research Center, from July 2006 to August 2007. Since September 2007, he has been with the Department of Electrical and Computer Engineering, McMaster University, where he is currently a Professor.

Dr. Chen held the title of the Barber-Gennum Chair in information technology, from 2008 to 2013, and the Joseph Ip Distinguished Engineering Fellow, from 2016 to 2018. He was a recipient of the Josef Raviv Memorial Postdoctoral Fellowship, in 2006, the Early Researcher Award from the Province of Ontario, in 2010, the IBM Faculty Award, in 2010, the JSPS Invitational Fellowship, in 2020, and the ICC Best Paper Award, in 2020. He served as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY, from 2014 to 2016. He is currently an Editor of the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING.

• • •