HIDeGan: A Hyperspectral-guided Image Dehazing GAN

Aditya Mehta, Harsh Sinha, Pratik Narang
Department of CSIS
BITS Pilani, India
{f2015808,h20130838,pratik.narang}@pilani.bits-pilani.ac.in

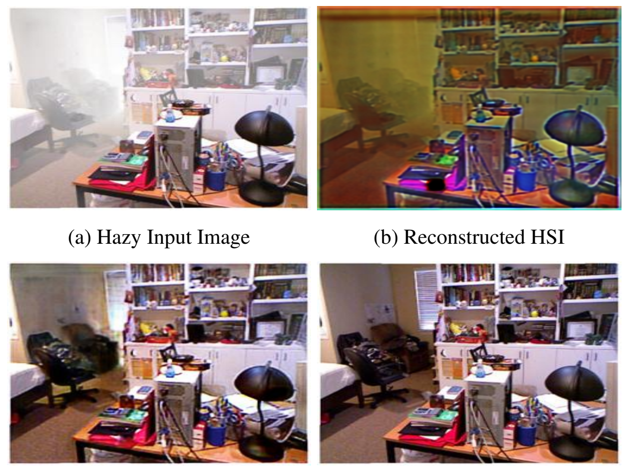
Murari Mandal
Department of CSE
MNIT Jaipur, India
murarimandal.cv@gmail.com

Abstract

*Haze removal in images captured from a diverse set of scenarios is a very challenging problem. The existing dehazing methods either reconstruct the transmission map or directly estimate the dehazed image in RGB color space. In this paper, we make a first attempt to propose a **H**yperspectral-guided **I**mage **De**hazing Generative Adversarial Network (*HIDeGan*). The* HIDeGan *architecture is formulated by designing a enhanced version of* CycleGAN *named* R2HCycle *and an enhanced conditional GAN named* H2RGan*. The* R2HCycle *makes use of the hyperspectral-image (HSI) in combination with cycle-consistency and skeleton losses in order to improve the quality of information recovery by analyzing the entire spectrum. The* H2RGan *estimates the clean RGB image from the hazy hyperspectral image generated by the* R2HCycle*. The models designed for spatial-spectral-spatial mapping generate visually better haze-free images. To facilitate HSI generation, datasets from spectral reconstruction challenge at NTIRE 2018 and NTIRE 2020 are used. A comprehensive set of experiments were conducted on the D-Hazy, HazezRD and the recent Reside-Standard (SOTS), Reside-β (SOTS) and Reside-β (HSTS) datasets. The proposed* HIDeGan *outperforms the existing state-of-the-art in all these datasets.*

1. Introduction

Downgraded visibility is typically the product of poor weather such as fog, snow, rain, and haziness in captured images. Haze refers to the deterioration in ambiance quality due to the changes in the concentration of particulate matter. Image quality is hampered by the variable density of particulate matter in the environment. Furthermore, under ambiguous conditions, floating objects such as darkness and smoke in the atmosphere consume and spread the light significantly and thus adversely affect image quality. Such visual disturbances also affect the performance of modern technologies based on vision such as object detection [1, 2, 3], segmentation [4, 5, 6, 7], object tracking [8, 9], etc. Thus, haze removal is an essential task for the proper functioning of several vision-based systems.



(a) Hazy Input Image (b) Reconstructed HSI

(c) Dehazed Output Image (d) Ground Truth

Figure 1: Example input hazy image, reconstructed hyperspectral image and the generated dehazed output image using HIDeGan. The effective dehazing and details of color and contrast can be compared to the ground truth image shown on bottom right.

The existing dehazing methods in the literature can be grouped into traditional and learning-based approaches [10, 11, 12]. In both the approaches, the physical scattering model [13] is frequently used to represent image formation. In this model, the image is formulated based on the properties of light transmission through the air. Most of the learning-based dehazing methods [14, 15] in the literature are based on the physical scattering model. The network usually learns the transmission map, which is converted into RGB image using the image formation model. However, the accuracy of the estimated atmospheric light and transmission map dramatically influences the quality of the dehazed image. The disjoint optimization of transmission map or atmospheric light may hamper the overall dehazing performance. Some recent approaches [16, 17] have directly

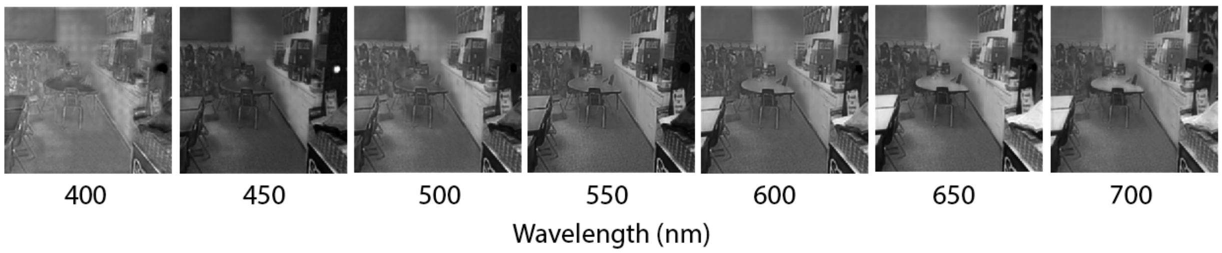


Figure 2: Comparison of reconstructed luminance across several spectral bands for a hazy input image. Visually, it can be inferred that different materials reflect and absorb differently. Spectral reflectance has a remarkable impact on image constrast. For instance, there is significant enhancement in tree-like texture (right of image) as we go to higher wavelengths.

estimated the clear image with RGB-RGB mapping

In this paper, we formulate the image dehazing problem as a hyperspectral image (HSI) guided image-to-image mapping task. Figure 1 shows a sample hazy image, the generated HSI, the dehazed image obtained using the proposed method, and its comparison to the ground truth. The proposed approach benefits from the spatial-spectral feature learning and is also free from the intermediate computation of transmission map. Our work is motivated by the study, analysis of HSI, and its effect on image quality [18]. HSI acquires spectral signatures from different wavelengths. As different materials reflect and absorb differently, the large pool of signals in HSI captured from different spectral channels are useful to discriminate between varieties of earth materials. The HSI facilitates the use of spatial relations between the various spectral responses in the vicinity, which is useful for better segmentation and classification of the image. Although the rich information can facilitate numerous applications, the storage requirements make it a very costly proposition. As a result, the use of HSI has been limited to preliminary analysis of observable signals in order to characterize the parts of the spectrum that carries valuable information for the application. Such applications include remote sensing [19, 20], astronomy [21], earth sciences[22], agriculture [23], and geology [24].

The use of HSI in general computer vision, in particular in the analysis of natural images, is still in its infancy. The main obstacles are lower resolution and higher cost of hyperspectral devices. Recently, researchers have focused on using approximation techniques which can reconstruct hyperspectral images from RGB images [25]. Researchers [26, 27] suggested a greater range of performance in the thermal IR band to look through the fog than in the visible band. Their models suggest that thermal imaging cameras can be useful for landing aids for aircraft or enhancing driver vision in the transportation and automotive industries. The models show that the fog penetration in the Long Wavelength Infrared (LWIR) is higher than the Medium Wavelength Infrared (MWIR) in all the cases tested. The results are corroborated by reconstructed hyperspectral images, as shown in Figure 2. In this paper, we try to extend this analysis as a significant incentive to analyze the entire spectrum for dehazing images. The main contributions of this paper can be summarized as follows:

1. We propose a hyperspectral guided generative adversarial network HIDEGAN for image dehazing. The HIDEGAN architecture is formulated by designing a CycleGAN named R2HCYCLE and a conditional GAN (cGAN) named H2RGAN. To the best of our knowledge, this is the first attempt to use HSI in GAN framework for haze removal.

2. The proposed R2HCYCLE makes use of the hyperspectral-image (HSI) in combination with cycle-consistency and skeleton losses in order to improve the quality of information recovery by analyzing the entire spectrum for dehazing. The H2RGAN generates the final dehazed output based on the RGB-HSI mapping learned from the R2HCYCLE.

3. HIDEGAN significantly outperforms the existing state-of-the-art methods in D-Hazy, HazeRD, and the more recent RESIDE-Standard (SOTS), RESIDE-β (SOTS) and RESIDE-β (HSTS) datasets. Furthermore, the detailed ablation study is carried out to analyze the effects of different components of the proposed network.

The rest of this paper is organized as follows. Section 2 discusses the related work. The proposed method is described in 3. Experimental results and discussion are delineated in Section 4, followed by the conclusion in 5.

2. Related Work

The dehazing methods are primarily aimed at restoring the clear image from a given hazy image. Existing techniques model scene reflection, atmospheric light and transmission map. The techniques can be divided into two broad

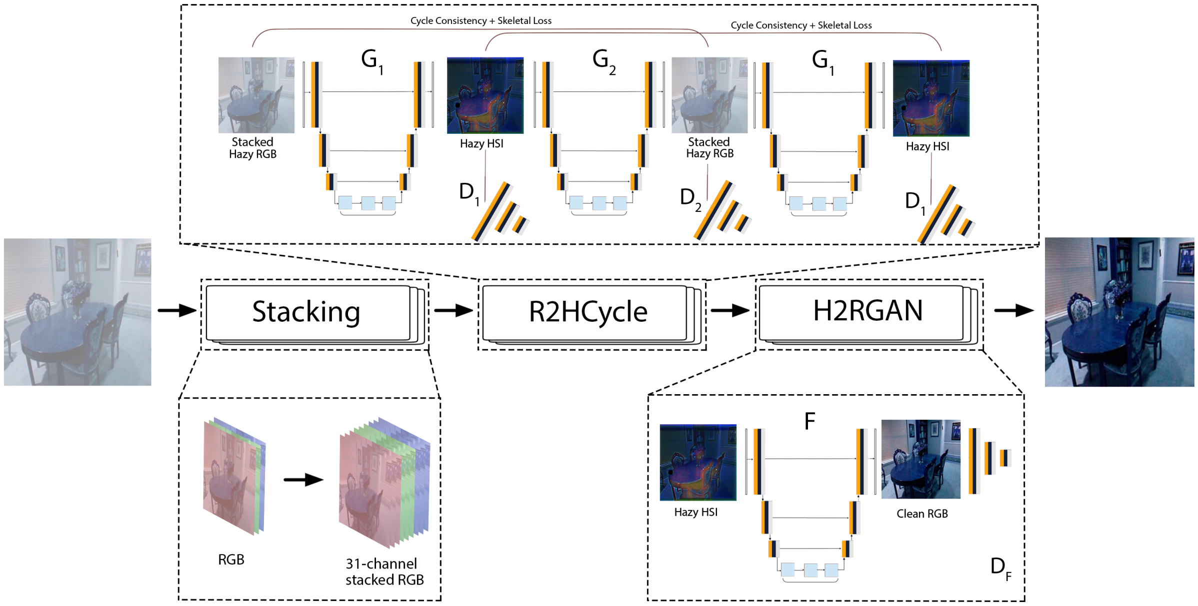


Figure 3: The architecture of HIDEGAN consists of two GANs, namely R2HCYCLE and H2RGAN. R2HCYCLE refers to GAN used for spectral reconstruction from hazy RGB images. The generated hyperspectral image is fed into H2RGAN to generate the corresponding clean RGB image. For, R2HCYCLE, G_1 & G_2 refers to the generators, and D_1 & D_2 to the discriminators. For the sake of clarity, the representation is split into two parts: hazy RGB to hazy HSI, and hazy HSI to hazy RGB image. For H2RGAN, F refers to the generator and D_F is the corresponding discriminator. Best viewed in color.

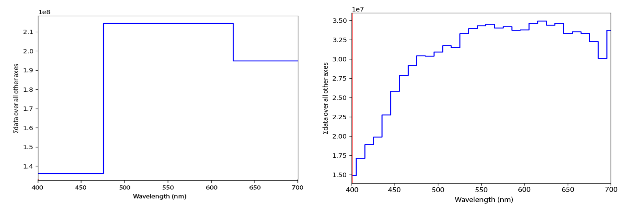
categories of solutions, namely, prior-based and learning-based methods. The prior-based methods can be further categorized in terms of multiple inputs images and polarizing filter based dehazing. The learning-based approaches adopt single image dehazing. These techniques also utilize additional information like depth or semantics.

Image dehazing was initially addressed by prior-based approaches [13]. He et al. [14] implemented dark channel prior (DCP) based on the statistics of clear images which help in estimating the transmission map. This was possible by utilizing dark pixels of different channels. For linear mapping of the local priors, Zhu et al. [28] proposed color attenuation prior (CAP). Used by Berman et al. [29], non-local color priors (NCP) model a hazy picture that incorporated a large number of distinct colors. Further, they employ clustering to represent a line in RGB space. Berman et al. [29] presented a haze-line prior-based approach to estimate the ambient atmospheric light. Similarly, a multi-scale approach was used for night-time dehazing by Ancuti et al. [15].

The dehazing task has also been tackled progressively by learning-based approaches. Such models typically CNNs and GANs are trained to learn the transmission map or atmospheric light [29]. Recently, deep-learning models have also been proposed for direct RGB-RGB mapping. Ren et al. proposed MSCNN [30] and GFN [31], where the former uses fine-scale local refinement holistic transmission map prediction, and the latter extracts multiple inputs that are further gated to dehaze the image. Several other CNN architectures [32, 33] were also presented in the literature for estimating the transmission maps. Cai et al. [10] proposed DehazeNet to estimate the intermediate transmission of map used to generate the haze-free image. Li et al. [34] engineered AOD-Net which has been able to produce dehazed images without any calculation for intermediate maps for learning a CNN dependent mapping feature for the reformulated physical scattering model.

Inspired by GAN's performance in image-to-image translation, several researchers have also solved the image dehazing problem by developing effective GAN architectures [16, 17]. Zhang et al. implemented DCPDN [35], which optimizes the relationship between hazy and haze-free image by estimating the transmission map and atmospheric light with two generators simultaneously. Existing GAN based techniques such as RI-GAN [11] and CD-Net [36] utilized cycle-consistency metrics while Yang et al. [37] tackled the problem by an unsupervised approach. Besides, Chen et al.[38] reformulated the concept by using a multiscale adaptive approach.

Hyperspectral images (HSI) were initially used for the



(a) Stacked RGB Image (b) Hyperspectral Image

Figure 4: The spectral response of an hyperspectral image (right) and its corresponding RGB image stacked to form 31 channels (left)

analysis of astronomical data [39] in spectrometers. Recently, researchers have focused on the reconstruction of HSI from RGB images using dictionaries, sparse coding, and manifold learning [25, 40]. The spectral response is, in practice, tough to learn, which makes the applicability of these approaches limited to specific scenarios. Therefore, CNN models have been proposed to overcome those problems [41, 42, 43, 44, 45]. Gwn Lore et al. [46] proposed a GAN based model to estimate the HSI from RGB input. Due to a lack of large-scale hyperspectral datasets, there has not been much focus on the use of hyperspectral guided approaches in image dehazing. Therefore, this work focuses on HSI reconstruction using an unsupervised approach. We apply and adapt the model to HSI reconstruction, which is, to the best of our knowledge, the first attempt for haze removal.

3. Proposed Method

The problem of reconstructing 31-channel hyperspectral images from 3-channel RGB input is clearly underconstrained. Hence, previous approaches have often relied on dictionary methods constructed from available large-scale hyperspectral image datasets. However, there are no datasets for hazy hyperspectral images.

We embrace the underlying uncertainty of the problem by posing it as an interpolation task. We stack 3-channel RGB image as a 31-channel image. Let an image matrix of height h and width w, consisting of 3 channels be represented as I_3, and a stacked image matrix consisting of 31 channels be represented as I_{31}.

$$
\begin{aligned}
I_3 &= (r_{h\times w\times 1} \mid g_{h\times w\times 1} \mid b_{h\times w\times 1}) \\
I_{31} &= (r_{h\times w\times 10} \mid g_{h\times w\times 10} \mid b_{h\times w\times 11})
\end{aligned} \tag{1}
$$

where r, g, b denote the red, green and blue channels respectively. Thus, the transformed problem is to interpolate the spectral response of stacked RGB input to an actual hyperspectral image as shown in Figure 4.

Hyperspectral Reconstruction [R2HCYCLE] We use an enhanced version of CycleGAN [47] to reconstruct hyperspectral images from stacked RGB images. The proposed architecture for interpolation, referred to as R2HCYCLE, is ultimately aimed at end-to-end hyperspectral reconstruction. It uses a cyclic skeleton-consistency loss in order to improve image consistency metrics apart from a min-max-based loss in the pixel domain. The essential idea behind cyclic skeleton-consistency loss is to compare the edges of the input image of the generated hyperspectral images. The proposed approach thus computes loss between the original image with the reconstructed hyperspectral image in both domains where the cyclic consistency ensures higher PSNR values, and the skeleton loss preserves the generated image sharpness.

For mapping function $G_1 : I_{31} \rightarrow X$ The adversarial loss for R2HCYCLE can be expressed as,

$$
\begin{aligned}
&\mathcal{L}_{GAN1}(G_1, D_1, I_{31}, X) \\
&= E_{i_{31}\sim p_{data(i_{31})}}[log(1 - D_1(G_1(i_{31})))] \\
&+ E_{x\sim p_{data(x)}}[log(D_1(x))]
\end{aligned} \tag{2}
$$

$$
\begin{aligned}
&\mathcal{L}_{GAN2}(G_2, D_2, X, I_{31}) \\
&= E_{x\sim p_{data(x)}}[log(1 - D_2(G_2(x)))] \\
&+ E_{i_{31}\sim p_{data(i_{31})}}[log(D_2(i_{31}))]
\end{aligned} \tag{3}
$$

$$
\mathcal{L}_{GAN} = \mathcal{L}_{GAN1} + \mathcal{L}_{GAN2} \tag{4}
$$

Cyclic skeleton-consistency loss Given, image $x \epsilon X$ domain & image $y \epsilon Y$ domain, a generator $G : X \rightarrow Y$& generator $F : Y \rightarrow X$, the formulation of cyclic skeleton-consistency loss is presented as,

$$
\begin{aligned}
\mathcal{L}_{skeleton} &= \|\psi(x) - \psi(F(G(x)))\|_2^2 \\
&+ \|\psi(y) - \psi(G(F(y)))\|_2^2.
\end{aligned} \tag{5}
$$

where (x, y) refers to stacked RGB image and hyperspectral image referred to as ground truth from unpaired image set. ψ denotes a Canny edge detector. The objective of proposed R2HCYCLE can be presented in Equation 6.

$$
\begin{aligned}
\mathcal{L}(G_1, G_2, D_1, D_2) &= \mathcal{L}_{GAN}(G_1, G_2, D_1, D_2) \\
&+ \lambda \cdot \mathcal{L}_{skeleton}(G_1, G_2)
\end{aligned} \tag{6}
$$

Image Dehazing [H2RGAN] Further, H2RGAN addresses the problem of estimating a clean RGB image from a hazy hyperspectral image. The reconstructed hyperspectral image is fed into an enhanced conditional GAN architecture, referred to as H2RGAN. The GAN is trained with reconstructed hyperspectral images as the input domain and clean ground truth RGB images as the output.

Let $x \; \epsilon \; \mathbb{R}^{h\times w\times 31}$ denote a reconstructed hyperspectral image and $y \; \epsilon \; \mathbb{R}^{h\times w\times 3}$ denote the corresponding clean image. The goal of H2RGAN is to seek a mapping that maps from x to y. It uses a min-max-based loss in the pixel domain for optimization of parameters.

We use the objective function of a conditional GAN which can be expressed as

$$\mathcal{L}_{cGAN}(F, D_F, X, Y, Z) = E_{x,z}[log(1 - D_F(x, F(x, z)))] + E_{x,y}[log(D(x, y))] \tag{7}$$

Network Architecture As demonstrated in Figure 3, the proposed architecture for R2HCYCLE consists of two generators G_1, G_2 and two discriminators D_1, D_2. For the sake of brevity, the figure shows both the generators in two different parts. In favor of generating a hyperspectral image, the architecture make use of cycle-consistency, cyclic skeleton-consistency losses and identity loss [47] besides the regular GAN discriminator and generator losses. As a result of this, the architecture is forced to preserve edge information of the input images and generate unique hyperspectral images. In addition, Figure 3 presents H2RGAN. It consists of a generator F and a discriminator D_F.

Implementation Details Zhu et al. [47] used a L1 distance in conjunction with the adversarial loss. As the problem is unconstrained in the case of R2HCYCLE, we explored L2 distance, which restricts the solution space. This ensures that the discriminator's task is to fool the generator while generating a conditional output close to the ground truth. We utilized L2 distance in R2HCYCLE as well as H2RGAN.

We adopt a U-Net with skip connections for generator architecture and PatchGAN for the discriminator. We utilize both bilinear interpolation layers and ConvTranspose2D layer for upsampling. All the components of HIDEGAN were implemented in PyTorch. We have used a batch size of 8, with Adam optimizer and a decaying learning rate. R2HCYCLE was trained for 20 epochs, while the H2RGAN ran for 50 epochs for both indoor and outdoor models. The training was carried out on datasets mentioned in Table 1 with NVIDIA RTX 2080ti GPU (11GB RAM).

4. Experimental Results and Discussion

In this section, we present a thorough evaluation of the proposed HIDEGAN. To determine the robustness of the proposed method, we examined it on both quantitative and qualitative grounds. For quantitative evaluation, metrics of structural similarity index (SSIM) [48] and peak signal to noise ratio (PSNR) are used, which are the most accepted and widely used metrics for dehazing algorithms assessment. Also, we conduct several ablation experiments in order to evaluate the contribution of various components in HIDEGAN.

Table 1: Dataset distribution for training and testing

Dataset	Training		Testing	
	Outdoor	Indoor	Outdoor	Indoor
RESIDE β OTS	14,000	—-	—-	—-
RESIDE Std ITS	—-	13,990	—-	—-
RESIDE HSTS	—-	—-	10	—-
RESIDE SOTS	—-	—-	500	500
HazeRD	75	—-	—-	—-
D-HAZY	—-	971	—-	478

Table 2: Comparative results of the proposed method and existing state-of-the-art dehazing methods over RESIDE-standard SOTS [49] Indoor dataset. (% inc denotes the percentage improvement for HIDEGAN over the given method)

Method	SSIM (%inc)	PSNR (% inc)
DCP [14]	0.8179 (6.15)	16.62 (48.67)
GRM [50]	0.8553 (1.51)	18.86 (31.01)
CDNet [36]	0.8852 (-1.92)	21.3 (16)
NLD [29]	0.7489 (15.93)	17.29 (42.91)
AOD-Net [34]	0.8504 (2.09)	19.06 (29.64)
DDN [37]	0.8242 (5.34)	19.38 (27.5)
CycleDehaze [16]	0.6923 (25.41)	15.86 (55.79)
GFN [31]	0.88 (-1.34)	22.31 (10.75)
C^2MSNet [51]	0.8152 (6.5)	20.12 (22.81)
RYFNet [12]	0.8716 (-0.39)	21.44 (15.25)
Pix2Pix [52]	0.82 (5.88)	16.84 (46.73)
RI-GAN [11]	0.85 (2.14)	19.828 (24.62)
HIDEGAN	**0.8682** (–)	**24.709 (–)**

Table 3: Comparative results of the proposed method and existing state-of-the-art dehazing methods over D-HAZY [15] dataset. (% inc denotes the percentage improvement for HIDEGAN over the given method)

Method	SSIM (%inc)	PSNR (% inc)
DCP [14]	0.706 (8.5)	11.59 (76.19)
DehazeNet [10]	0.727 (5.36)	13.4 (52.39)
CDNet [36]	0.7411 (3.36)	13.84 (47.54)
C^2MSNet [51]	0.7201 (6.37)	12.71 (60.66)
MSCNN [30]	0.7231 (5.93)	12.82 (59.28)
AODNet [34]	0.7177 (6.73)	12.41 (64.54)
Pix2Pix [52]	0.7519 (1.88)	16.43 (24.28)
RI-GAN [11]	0.8179 (-6.35)	18.8167 (8.52)
DDN [32]	0.7383 (3.75)	10.96 (86.31)
CycleGAN [47]	0.649 (18.03)	13.69 (49.16)
CycleDehaze [16]	0.6746 (13.55)	12.54 (62.84)
HIDEGAN	**0.766 (–)**	**20.42 (–)**

4.1. Datasets

The ICVL BGU Hyperspectral dataset [25] is used in training R2HCycle for hyperspectral reconstruction. It consists of 200 natural images, consisting of houses, trees, specific indoor scenes, etc. Every hyperspectral image has a size of 1392 $\times$1300 with 519 bands each. However, the dataset also provides sampled images consisting of 31 bands separated by roughly 10nm each to address computational constraints. The same dataset was used for hyperspectral reconstruction challenge as a part of NTIRE 2018 [53]. In addition, we also used hyperspectral images from NTIRE 2020 hyperspectral reconstruction challenge which consists of 360 hyperspectral images. As a part of data augmentation, like flipping and random cropping were carried out to obtain 11000 hyperspectral images.

The quantitative comparisons were made using four datasets – RESIDE-SOTS indoor dataset (500 images), RESIDE-SOTS outdoor dataset (500 images), RESIDE HSTS (10 images) [49], and D-Hazy (1499 images) [15]. RESIDE is among the benchmark datasets for image dehazing, and it also provides benchmarking of nine representative state-of-the-art dehazing networks by offering full reference evaluation metrics such as PSNR and SSIM for the synthetic objective test set (SOTS). The RESIDE-hybrid subjective test collection (HSTS) offers 10 hazy outdoor synthetic images for test purposes only. The HazeRD dataset [54] is composed of 15 outdoor real-world scenes. Five different weather conditions are simulated for each scene which lead to 75 pairs of hazy and clean images. Similarly, the D-Hazy dataset [15] includes 1449 pairs of clear and synthesized hazy images that can be used to evaluate haze removal.

We have used a large variety of datasets for training and testing purposes to ensure that HIDEGAN does not overfit on any single dataset, and effectively eliminates haze by learning the underlying task. It also means that HIDEGAN is able to work under varying conditions of haze. The detailed distribution of the data set for the purposes of training and testing is given in Table 1.

4.2. Quantitative results

We report the average PSNR and SSIM of all stated networks and the proposed method. Along with the respective values, we also report the percentage increase (in SSIM or PSNR) achieved by our model as compared to each method. According to Tables 2, 3, 4 and 5, HIDEGAN achieves impressive performance and outperforms all the approaches in terms of PSNR. We also achieve superior or comparable results in terms of SSIM.

RESIDE dataset (SOTS indoor/outdoor, HSTS). Table 2 displays the comparative findings collected over 17 state-of-the-art methods for the RESIDE-SOTS indoor dataset along with the proposed HIDEGAN. The results clearly indicate that the proposed model is much more effective than others, especially in terms of PSNR. The improvement margin in PSNR ranges from 10.75 to 74.5 percent. In comparison with all 7 current state-of-the-art methods benchmarked over the RESIDE-SOTS outdoor images (Table 4), we witness a similar improvement in both SSIM and PSNR. Even from Table 5, we can observe a significant improvement in performance in terms of PSNR over the HSTS dataset.

Table 4: Comparative results of the proposed method and existing state-of-the-art dehazing methods over RESIDE-standard SOTS [49] Outdoor dataset. (% inc denotes the percentage improvement for HIDEGAN over the given method)

Method	SSIM (%inc)	PSNR (% inc)
DCP [14]	0.8148 (7.76)	19.13 (33.49)
DehazeNet [10]	0.8514 (3.12)	22.46 (13.7)
AOD-Net [34]	0.8765 (0.17)	20.29 (25.86)
MADN [58]	0.9137 (-3.91)	23.64 (8.02)
GFN [31]	0.8444 (3.98)	21.55 (18.5)
Enh. Pix2Pix [17]	0.863 (1.74)	22.57 (13.14)
HIDEGAN	**0.878 (–)**	**25.536 (–)**

D-Hazy dataset. D-Hazy dataset was evaluated on the indoor model. This dataset contains images of relatively denser haze. Nonetheless, as is evident form Table 3, the proposed method outperforms existing solutions. The improvement in PSNR ranges from 8.52 to 76.19 percent. The improvement in SSIM is also substantial.

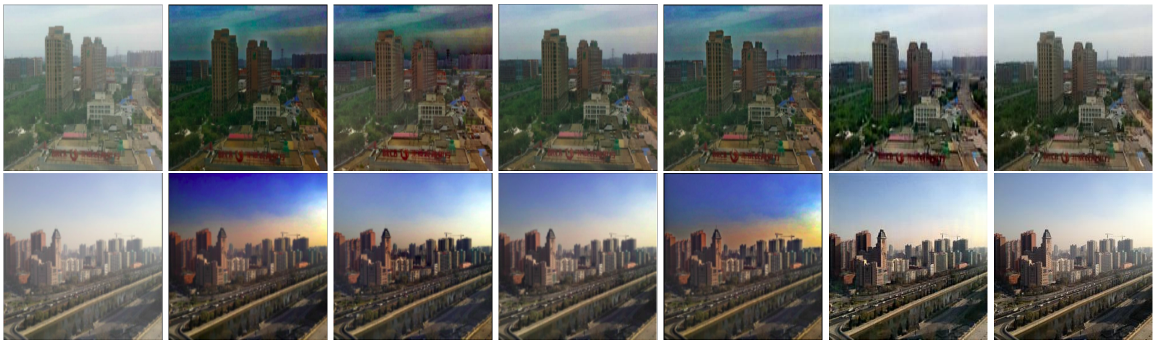
4.3. Qualitative results

The efficacy of the hyperspectral-guided approach of HIDEGAN is demonstrated by a qualitative analysis and comparison with existing approaches.

Figure 5 compares the performance of the proposed approach with several prior based approaches and compares them to the ground truth (denoted as GT). It can be observed that prior based approaches are not able to retain the original color and contrast of the image, while HIDEGAN renders these details faithfully and its dehazed output is the closest to the ground truth.

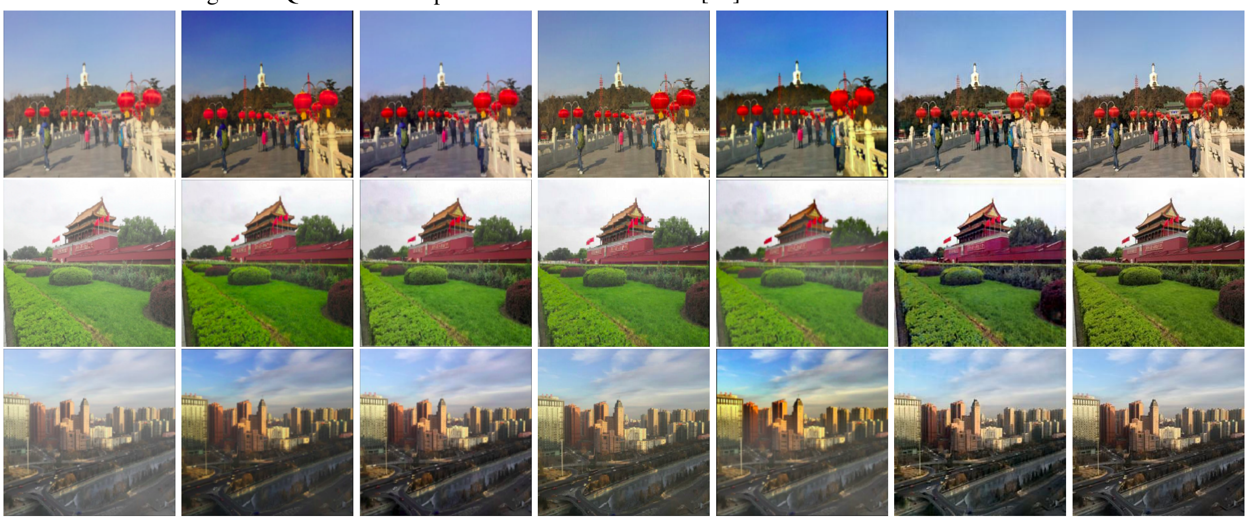
In Figure 6, HIDEGAN is compared with learning based approaches (the ground truth is depicted as GT in the image). It can be observed that dehazed output DeepDCP [57] still has a hazy tinge while the original colors have been distorted in MSCNN [30]. HIDEGAN also achieves sharper details and edges as compared to other approaches.

Finally, we evaluate the performance of HIDEGAN on real-world hazy images (for which the ground truth is not available) as shown in Figure 7. In comparison to the exist-



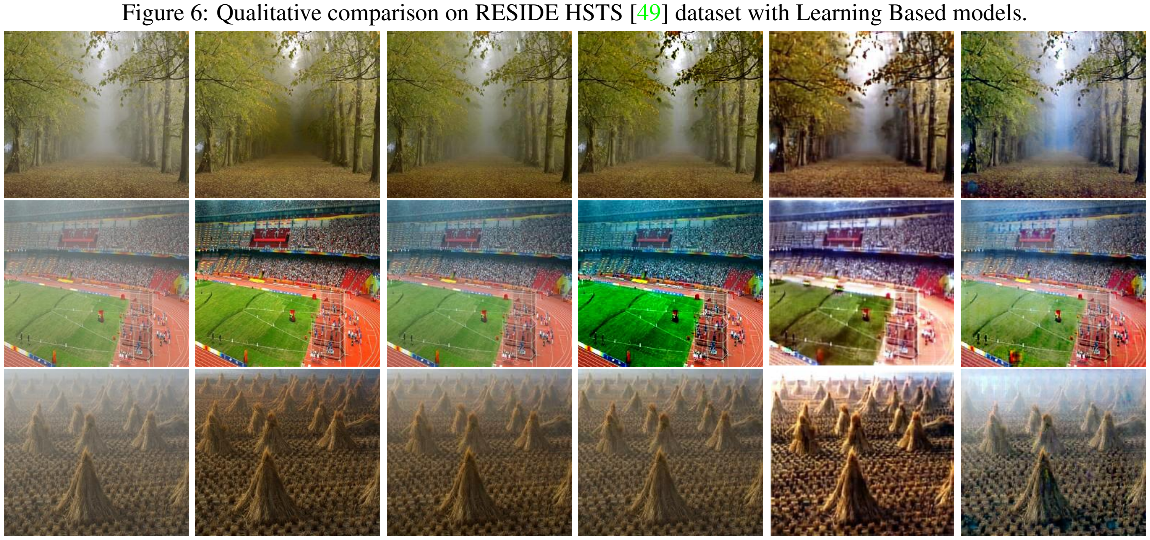
(a) Hazy Input (b) DCP [14] (c) NLD [29] (d) CAP [55] (e) BCCR [56] (f) **HIDeGan** (g) GT

Figure 5: Qualitative comparision on RESIDE HSTS [49] dataset with Prior Based models.



(a) Hazy Input (b) AOD [34] (c) DehazeNet [10] (d) DeepDCP [57] (e) MSCNN [30] (f) **HIDeGan** (g) GT

Figure 6: Qualitative comparison on RESIDE HSTS [49] dataset with Learning Based models.



(a) Hazy Input (b) DCP [14] (c) DehazeNet [10] (d) MSCNN [30] (e) CycleDehaze [16] (f) **HIDeGan**

Figure 7: Qualitative comparison on natural hazy images with state-of-the-art-results

Table 5: Comparative results of the proposed method and existing state-of-the-art dehazing methods over HSTS [49] dataset. (% inc denotes the percentage improvement for HIDEGAN over the given method)

Method	SSIM (%inc)	PSNR (% inc)
DCP [14]	0.7609 (17.49)	14.84 (88.96)
FVR [59]	0.7624 (17.26)	14.48 (93.66)
BCCR [56]	0.7382 (21.11)	15.08 (85.95)
GRM [50]	0.8184 (9.24)	18.54 (51.25)
Deep DCP [57]	0.933 (-4.18)	24.44 (14.74)
NLD [29]	0.7411 (20.63)	18.92 (48.21)
DehazeNet [10]	0.9153 (-2.33)	24.48 (14.55)
MSCNN [30]	0.8168 (9.45)	18.64 (50.44)
AODNet [34]	0.8973 (-0.37)	20.55 (36.46)
GFN [31]	0.874 (2.29)	22.94 (22.24)
HIDEGAN	**0.894 (–)**	**28.042 (–)**

ing methods, HIDEGAN recovers uniform color and contrast details. It can be clearly seen that images dehazed by HIDEGAN have less noise and sharper edges for both synthetic and real-world hazy images.

4.4. Ablation study

In this subsection, we perform a component-wise ablation analysis to understand the contribution of the major components of our network towards effective dehazing. We have considered 4 baseline models apart from the final model, HIDEGAN. Two of the four baseline works are CycleGAN [47] and Pix2Pix [52]. The performances of all models are compared on RESIDE-SOTS [49] indoor dataset and summarized in Table 6.

Hyper-spectral imaging. The GAN architectures of CycleGAN [47] and pix2pix [52], which perform dehazing by using the RGB images, do not fare well as compared to HIDEGAN, which can be directly observed in terms of the PSNR and SSIM reported in Table 6. The performance of HIDEGAN clearly shows the benefit of using HSI information over RGB. This component is represented by "HSI" column name in the Table 6 (✓means hyperspectral images were used in the model).

Stacked RGB input (I_{31}) for generating HSI images using R2HCYCLE. This part is referenced in Table 6 as I_{31}, and we consider Model 1 without the stacked RGB input. It can be easily inferred from the results of Model 1 that HSI generation, being an ill-posed problem, becomes worse when it has to be performed only from three channels. Since the intermediate HSI obtained is poor in that case, H2RGAN also fails to fare well. Stacking the RGB to R2HCYCLE with a 31 channel input clearly enhances its performance and thus greatly improves the output.

Effect of loss. The effect of skeleton loss in Eq. 5 is verified in Table 6. Model 2 does not incorporate skeleton loss. The structure of objects present in an image is one of the typical underlying detail which remains intact in an RGB image and its corresponding hyper-spectral image. Thus, adding a loss that penalizes the distorted edges helps the network learn better. This is confirmed by the contrast between the results obtained with Model 2 and HIDEGAN.

Table 6: Component-wise Ablation Study

Model	Components			SOTS	
	HSI	Skeleton Loss	I_{31}	SSIM	PSNR
CycleGAN[47]				0.5738	14.16
pix2pix[52]				0.82	16.84
Model 1	✓	✓		0.486	13.17
Model 2	✓		✓	0.8279	21.83
HIDEGAN	✓	✓	✓	**0.8682**	**24.709**

The component-wise ablation study conclusively demonstrates that HIDEGAN, utilizing hyperspectral information, I_{31} input and skeleton loss is superior to models not having any one of them.

5. Conclusion

This paper presented a first attempt at a hyperspectral-guided generative adversarial network HIDEGAN for image dehazing. The HIDEGAN architecture uses an enhanced cycle GAN (R2HCYCLE) which utilized HSI in combination with cycle-consistency and skeleton losses in order to improve the quality of information recovery by analyzing the entire spectrum for dehazing. The architecture further used an enhanced conditional GAN (H2RGAN) which generates the final dehazed image based on the RGB-HSI mapping learned from the R2HCYCLE. A detailed ablation study demonstrated the worth of the individual components of HIDEGAN. HIDEGAN also outperforms the existing state-of-the-art methods quantitatively and qualitatively.

Acknowledgement

The authors are thankful to NVIDIA for providing TITAN Xp GPU grant. We would also like to thank Mr Aryan Mehra for his valuable support.

References

[1] Linghua Zhou, Weidong Min, Deyu Lin, Qing Han, and Ruikang Liu, "Detecting motion blurred vehicle logo in iov using filter-deblurgan and vl-yolo," *IEEE Transactions on Vehicular Technology*, 2020. 1

[2] Murari Mandal, Manal Shah, Prashant Meena, Sanhita Devi, and Santosh Kumar Vipparthi, "Avdnet: A small-sized ve-

hicle detection network for aerial visual data," *IEEE Geoscience and Remote Sensing Letters*, 2019. 1

[3] Murari Mandal, Manal Shah, Prashant Meena, and Santosh Kumar Vipparthi, "Sssdet: Simple short and shallow network for resource efficient vehicle detection in aerial scenes," in *2019 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2019, pp. 3098–3102. 1

[4] Murari Mandal, Vansh Dhar, Abhishek Mishra, and Santosh Kumar Vipparthi, "3dfr: A swift 3d feature reductionist framework for scene independent change detection," *IEEE Signal Processing Letters*, vol. 26, no. 12, pp. 1882–1886, 2019. 1

[5] Murari Mandal, Mallika Chaudhary, Santosh Kumar Vipparthi, Subrahmanyam Murala, Anil Balaji Gonde, and Shyam Krishna Nagar, "Antic: Antithetic isomeric cluster patterns for medical image retrieval and change detection," *IET Computer Vision*, vol. 13, no. 1, pp. 31–43, 2018. 1

[6] Murari Mandal, Prafulla Saxena, Santosh Kumar Vipparthi, and Subrahmanyam Murala, "Candid: Robust change dynamics and deterministic update policy for dynamic background subtraction," in *2018 24th International Conference on Pattern Recognition (ICPR)*. IEEE, 2018, pp. 2468–2473. 1

[7] Thangarajah Akilan, QM Jonathan Wu, and Wandong Zhang, "Video foreground extraction using multi-view receptive field and encoder–decoder dcnn for traffic and surveillance applications," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9478–9493, 2019. 1

[8] Pei-Hsuan Chiu, Po-Hsuan Tseng, and Kai-Ten Feng, "Interactive mobile augmented reality system for image and hand motion tracking," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9995–10009, 2018. 1

[9] Murari Mandal, Lav Kush Kumar, Mahipal Singh Saran, and Santosh Kumar vipparthi, "Motionrec: A unified deep framework for moving object recognition," in *The IEEE Winter Conference on Applications of Computer Vision (WACV)*, March 2020. 1

[10] 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*, vol. 25, no. 11, pp. 5187–5198, 2016. 1, 3, 5, 6, 7, 8

[11] Akshay Dudhane, Harshjeet Singh Aulakh, and Subrahmanyam Murala, "Ri-gan: An end-to-end network for single image haze removal," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 0–0. 1, 3, 5

[12] Akshay Dudhane and Subrahmanyam Murala, "Ryf-net: Deep fusion network for single image haze removal," *IEEE Transactions on Image Processing*, 2019. 1, 5

[13] Robby T Tan, "Visibility in bad weather from a single image," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2008, pp. 1–8. 1, 3

[14] Kaiming He, Jian Sun, and Xiaoou Tang, "Single image haze removal using dark channel prior," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 12, pp. 2341–2353, 2010. 1, 3, 5, 6, 7, 8

[15] Cosmin Ancuti, Codruta O Ancuti, and Christophe De Vleeschouwer, "D-hazy: A dataset to evaluate quantitatively dehazing algorithms," in *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2016, pp. 2226–2230. 1, 3, 5, 6

[16] Deniz Engin, Anil Genç, and Hazim Kemal Ekenel, "Cycle-dehaze: Enhanced cyclegan for single image dehazing," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 825–833. 1, 3, 5, 7

[17] Yanyun Qu, Yizi Chen, Jingying Huang, and Yuan Xie, "Enhanced pix2pix dehazing network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 8160–8168. 1, 3, 6

[18] John P Kerekes and John R Schott, "Hyperspectral imaging systems," *Hyperspectral data exploitation: Theory and applications*, pp. 19–45, 2007. 2

[19] Thomas Lillesand, Ralph W Kiefer, and Jonathan Chipman, *Remote sensing and image interpretation*, John Wiley & Sons, 2015. 2

[20] Shivangi Dwivedi, Murari Mandal, Shekhar Yadav, and Santosh Kumar Vipparthi, "3d cnn with localized residual connections for hyperspectral image classification," in *Computer Vision and Image Processing: 4th International Conference, CVIP 2019, Jaipur, India, September 27–29, 2019, Revised Selected Papers, Part II 4*. Springer, 2020, pp. 354–363. 2

[21] E Keith Hege, Dan O'Connell, William Johnson, Shridhar Basty, and Eustace L Dereniak, "Hyperspectral imaging for astronomy and space surveillance," in *Imaging Spectrometry IX*. International Society for Optics and Photonics, 2004, vol. 5159, pp. 380–391. 2

[22] John F Mustard and Jessica M Sunshine, "Spectral analysis for earth science: investigations using remote sensing data," *Remote sensing for the earth sciences: Manual of remote sensing*, vol. 3, pp. 251–307, 1999. 2

[23] Driss Haboudane, John R Miller, Elizabeth Pattey, Pablo J Zarco-Tejada, and Ian B Strachan, "Hyperspectral vegetation indices and novel algorithms for predicting green lai of crop canopies: Modeling and validation in the context of precision agriculture," *Remote sensing of environment*, vol. 90, no. 3, pp. 337–352, 2004. 2

[24] Edward A Cloutis, "Review article hyperspectral geological remote sensing: evaluation of analytical techniques," *International Journal of Remote Sensing*, vol. 17, no. 12, pp. 2215–2242, 1996. 2

[25] Boaz Arad and Ohad Ben-Shahar, "Sparse recovery of hyperspectral signal from natural rgb images," in *European Conference on Computer Vision*. Springer, 2016, pp. 19–34. 2, 4, 6

[26] Victor J Norris, Robert S Evans, and Douglas G Currie, "Performance comparison of visual, infrared, and ultraviolet sensors for landing aircraft in fog," in *Gateway to the New Millennium. 18th Digital Avionics Systems Conference. Proceedings (Cat. No. 99CH37033)*. IEEE, 1999, vol. 1, pp. 5–D. 2

[27] Bin Wang, Lili Dong, Ming Zhao, Houde Wu, Yuanyuan Ji, and Wenhai Xu, "An infrared maritime target detection algorithm applicable to heavy sea fog," *Infrared Physics & Technology*, vol. 71, pp. 56–62, 2015. 2

[28] Hongyuan Zhu, Xi Peng, Vijay Chandrasekhar, Liyuan Li, and Joo-Hwee Lim, "Dehazegan: When image dehazing meets differential programming.," in *IJCAI*, 2018, pp. 1234–1240. 3

[29] Dana Berman, Shai Avidan, et al., "Non-local image dehazing," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 1674–1682. 3, 5, 7, 8

[30] Wenqi Ren, Si Liu, Hua Zhang, Jinshan Pan, Xiaochun Cao, and Ming-Hsuan Yang, "Single image dehazing via multi-scale convolutional neural networks," in *European conference on computer vision*. Springer, 2016, pp. 154–169. 3, 5, 6, 7, 8

[31] Wenqi Ren, Lin Ma, Jiawei Zhang, Jinshan Pan, Xiaochun Cao, Wei Liu, and Ming-Hsuan Yang, "Gated fusion network for single image dehazing," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 3253–3261. 3, 5, 6, 8

[32] Xi Yang, Hui Li, Yu-Long Fan, and Rong Chen, "Single image haze removal via region detection network," *IEEE Transactions on Multimedia*, 2019. 3, 5

[33] Dong Yang and Jian Sun, "Proximal dehaze-net: a prior learning-based deep network for single image dehazing," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 702–717. 3

[34] 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*, 2017, pp. 4770–4778. 3, 5, 6, 7, 8

[35] He Zhang and Vishal M Patel, "Densely connected pyramid dehazing network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 3194–3203. 3

[36] Akshay Dudhane and Subrahmanyam Murala, "Cdnet: Single image de-hazing using unpaired adversarial training," in *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2019, pp. 1147–1155. 3, 5

[37] Xitong Yang, Zheng Xu, and Jiebo Luo, "Towards perceptual image dehazing by physics-based disentanglement and adversarial training," in *Thirty-second AAAI conference on artificial intelligence*, 2018. 3, 5

[38] Shuxin Chen, Yizi Chen, Yanyun Qu, Jingying Huang, and Ming Hong, "Multi-scale adaptive dehazing network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 0–0. 3

[39] Robert O. Green, Betina E. Pavri, and Thomas G. Chrien, "On-orbit radiometric and spectral calibration characteristics of EO-1 hyperion derived with an underflight of AVIRIS and in situ measurements at salar de arizaro, argentina," *IEEE Trans. Geoscience and Remote Sensing*, vol. 41, no. 6, pp. 1194–1203, 2003. 4

[40] Yan Jia, Yinqiang Zheng, Lin Gu, Art Subpa-Asa, Antony Lam, Yoichi Sato, and Imari Sato, "From rgb to spectrum for natural scenes via manifold-based mapping," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 4705–4713. 4

[41] Zhiwei Xiong, Zhan Shi, Huiqun Li, Lizhi Wang, Dong Liu, and Feng Wu, "Hscnn: Cnn-based hyperspectral image recovery from spectrally undersampled projections," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 518–525. 4

[42] Silvano Galliani, Charis Lanaras, Dimitrios Marmanis, Emmanuel Baltsavias, and Konrad Schindler, "Learned spectral super-resolution," *arXiv preprint arXiv:1703.09470*, 2017. 4

[43] Yigit Baran Can and Radu Timofte, "An efficient cnn for spectral reconstruction from rgb images," *arXiv preprint arXiv:1804.04647*, 2018. 4

[44] Z Shi, C Chen, Z Xiong, D Liu, and F Hscnn Wu, "Advanced cnn-based hyperspectral recovery from rgb images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA*, 2018, pp. 18–22. 4

[45] Tarek Stiebel, Simon Koppers, Philipp Seltsam, and Dorit Merhof, "Reconstructing spectral images from rgb-images using a convolutional neural network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 948–953. 4

[46] Kin Gwn Lore, Kishore K Reddy, Michael Giering, and Edgar A Bernal, "Generative adversarial networks for spectral super-resolution and bidirectional rgb-to-multispectral mapping," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 0–0. 4

[47] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232. 4, 5, 8

[48] Alain Hore and Djemel Ziou, "Image quality metrics: Psnr vs. ssim," in *2010 20th International Conference on Pattern Recognition*. IEEE, 2010, pp. 2366–2369. 5

[49] 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*, vol. 28, no. 1, pp. 492–505, 2018. 5, 6, 7, 8

[50] Chen Chen, Minh N Do, and Jue Wang, "Robust image and video dehazing with visual artifact suppression via gradient residual minimization," in *European Conference on Computer Vision*. Springer, 2016, pp. 576–591. 5, 8

[51] Akshay Dudhane and Subrahmanyam Murala, "Cˆ 2msnet: A novel approach for single image haze removal," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2018, pp. 1397–1404. 5

[52] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134. 5, 8

[53] Boaz Arad, Ohad Ben-Shahar, and Radu Timofte, "Ntire 2018 challenge on spectral reconstruction from rgb images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 929–938. 6

[54] Yanfu Zhang, Li Ding, and Gaurav Sharma, "Hazerd: an outdoor scene dataset and benchmark for single image dehazing," in *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2017, pp. 3205–3209. 6

[55] Qingsong Zhu, Jiaming Mai, and Ling Shao, "A fast single image haze removal algorithm using color attenuation prior," *IEEE transactions on image processing*, vol. 24, no. 11, pp. 3522–3533, 2015. 7

[56] Gaofeng Meng, Ying Wang, Jiangyong Duan, Shiming Xiang, and Chunhong Pan, "Efficient image dehazing with boundary constraint and contextual regularization," in *Proceedings of the IEEE international conference on computer vision*, 2013, pp. 617–624. 7, 8

[57] Alona Golts, Daniel Freedman, and Michael Elad, "Unsupervised single image dehazing using dark channel prior loss," *arXiv preprint arXiv:1812.07051*, 2018. 6, 7, 8

[58] Shuxin Chen, Yizi Chen, Yanyun Qu, Jingying Huang, and Ming Hong, "Multi-scale adaptive dehazing network," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2019. 6

[59] Jean-Philippe Tarel and Nicolas Hautiere, "Fast visibility restoration from a single color or gray level image," in *2009 IEEE 12th International Conference on Computer Vision*. IEEE, 2009, pp. 2201–2208. 8