

Domain-Aware Unsupervised Hyperspectral Reconstruction for Aerial Image Dehazing

Aditya Mehta, Harsh Sinha, Pratik Narang
Department of CSIS
BITS Pilani, India

Murari Mandal
Department of CSE
MNIT Jaipur, India

{f2015808,h20130838,pratik.narang}@pilani.bits-pilani.ac.in

murarimandal.cv@gmail.com

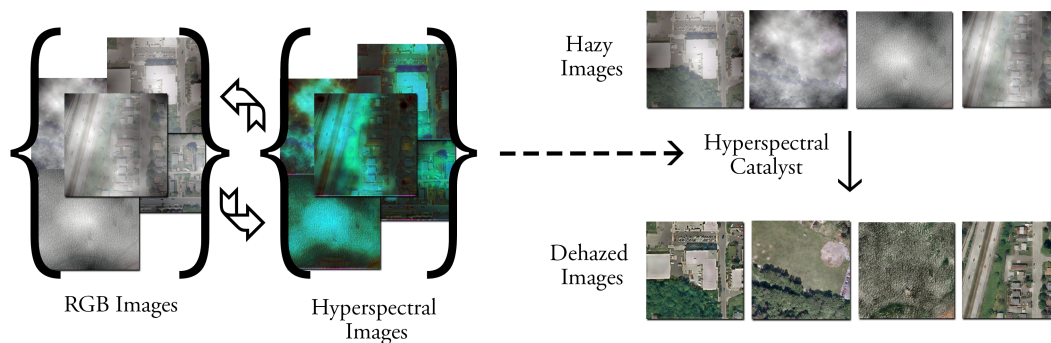


Figure 1: Left: Unsupervised hyperspectral reconstruction for hazy aerial images. Right: Once trained, features derived from reconstructed hyperspectral images serve as a catalyst to dehaze aerial images

Abstract

Haze removal in aerial images is a challenging problem due to considerable variation in spatial details and varying contrast. Changes in particulate matter density often lead to degradation in visibility. Therefore, several approaches utilize multi-spectral data as auxiliary information for haze removal. In this paper, we propose SkyGAN for haze removal in aerial images. SkyGAN consists of 1) a domain-aware hazy-to-hyperspectral (H2H) module, and 2) a conditional GAN (cGAN) based multi-cue image-to-image translation module (I2I) for dehazing. The proposed H2H module reconstructs several visual bands from RGB images in an unsupervised manner, which overcomes the lack of hazy hyperspectral aerial image datasets. The module utilizes task supervision and domain adaptation in order to create a “hyperspectral catalyst” for image dehazing. The I2I module uses the hyperspectral catalyst along with a 12-channel multi-cue input and performs effective image dehazing by utilizing the entire visual spectrum. In addition, this work introduces a new dataset, called Hazy Aerial-Image (HAI) dataset, that contains more than 65,000 pairs of hazy and ground truth aerial images with realistic, non-

homogeneous haze of varying density. The performance of SkyGAN is evaluated on the recent SateHaze1k dataset as well as the HAI dataset. We also present a comprehensive evaluation of HAI dataset with a representative set of state-of-the-art techniques in terms of PSNR and SSIM.

1. Introduction

Aerial imagery refers to photographs taken from aircraft such as helicopters and UAVs. Such images are advantageous as they have rich information content. Hence, they have been widely utilized in various fields such as remote sensing [21], earth sciences [33], agriculture [15] and geology [6]. The aerial images and video data facilitate numerous applications such as aerial surveillance [27, 28, 29], search and rescue, event recognition [32], urban and rural scene understanding [43]. As the aerial images are perceived from a considerable distance, these images often suffer from low visibility, color shift, and blurriness due to changes in the atmospheric path. Such atmospheric effects, especially non-homogeneous clouds, fog, and haze, degrade the quality of input images. Therefore, there is a need to address visibility improvement before aerial images can be



Figure 2: Illustration of corresponding clean and hazy images from the proposed HAI dataset which consists of 5 different degrees of haze, represented in different columns.

used in aerial vision-based systems.

Researchers have used several different techniques to remove haze in aerial images. Zhang et al. [48] used a correction technique using a correlation between the high-frequency and the low-frequency color bands. Such methods quantify the amount of haze using different spectral bands in the visible region. Liu et al. [22] present a virtual cloud point method based on relative haze thickness for haze removal. Long et al. [24] removed haze in remote sensing images by adapting the DCP proposed by He et al. [16] for natural scene images. Similarly, Shen et al. [38] a gradient-based spectral adaptive approach to exploit the wavelength-dependent transmission information. Makarau et al. [26] have used NIR and multispectral images to exploit the multi-sensor information for accurate haze removal. However, the use of hyperspectral images (HSI) in deep learning-based haze removal has been limited. HSI consists of spectral reflectance information from a substantial number of wavebands. The high cost of HSI acquisition devices and the lack of large-scale HSI datasets are some of the impediments in wide-scale acceptance of HSI in general computer vision. Acquisition of full spectral signatures using a HSI camera is a costly endeavor, not only in terms of hardware but it also hampers the viability of UAVs and other radio-controlled aerial vehicles.

In this work, the input hazy RGB image is used to reconstruct a domain-aware HSI using a modified cycle-consistent framework. The proposed framework is a low-cost and efficient approach to generate HSI in an unsupervised manner. Further, we present the practical applicability of reconstructed HSI for aerial image dehazing. The key motivation to use HSI for dehazing is to utilize the entire visual spectrum which provides rich information content in comparison to three primary bands: red, green and blue. Since the perception of an object depends on its nature to absorb different wavelengths, this work proposes to enhance the visual contrast of an image by acquiring infor-

mation from different spectral regions.

The key contributions of this work are as follows:

1. A novel GAN framework named SkyGAN is presented which incorporates HSI guidance in image-to-image translation network for image dehazing. To the best of our knowledge, it is the first attempt to use HSI in a GAN framework for aerial image dehazing.
2. SkyGAN reconstructs HSI from RGB images in an unpaired manner, which addresses the lack of hazy HSI datasets.
3. The architecture utilizes task supervision in conjunction with learning closed set domain adaptation by combining two techniques, adversarial distribution discrepancy alignment [10] and cycle-consistency constraint [50, 45, 18], to reconstruct domain-aware HSI from a set of natural scene HSI.
4. A residual network is used to alleviate data distribution discrepancy from reconstructed HSI, thereby generating a hyperspectral catalyst (HSC) which is fed along with a 12-channel multi-cue input to the proposed conditional GAN (cGAN) for haze removal.
5. Further, we introduce a large-scale hazy aerial image dataset HAI, consisting of over 65 thousand pairs of hazy and ground truth (haze-free) aerial images. The images in HAI contain realistic, non-homogeneous haze of varying density.
6. The proposed SkyGAN outperforms the existing state-of-the-art approaches on the benchmark SateHaze1k [17] and the new large-scale HAI dataset. We also present an extensive ablation study of the different components of SkyGAN.

2. Related Work

The existing dehazing approaches can be categorized into radiative transfer (RT) models which require atmosphere ambience related parametric information, and statistical information (SI) methods which rely on image char-

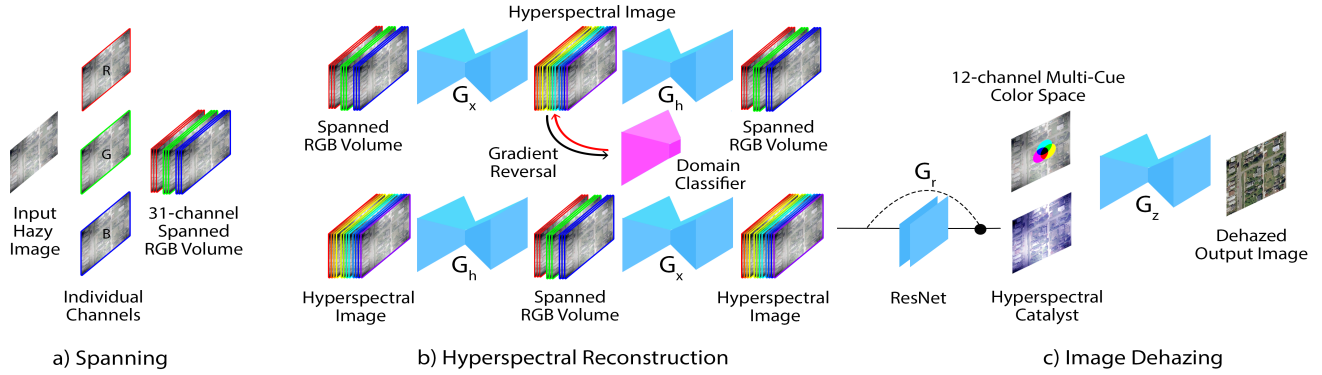


Figure 3: Architecture for the proposed SkyGAN. The corresponding discriminators D_x, D_h and D_z for generators G_x, G_h and G_z respectively are not displayed for brevity of the diagram.

acteristics for haze removal. RT models build upon some assumption or a physical prior such as contrast [40], dark channel [16], color-attenuation [52], and non-local prior [4]. On the other hand, there have been several SI approaches that typically learn the transmission map [5] or a direct end-to-end mapping between hazy and clean images [19, 34]. With advancements in deep learning, researchers have employed GANs for image dehazing task by learning the transmission map [46, 8, 35, 7]. Researchers have also used hyperspectral images with GANs for super-resolution [14], classification [51] and natural scene dehazing [30].

The existing techniques use standard RGB images for dehazing. HSI can be utilized to enhance segmentation and classification by using rich information from the entire visual spectrum. The use of HSI has been restricted primarily due to its high acquisition cost. Therefore researchers have developed techniques such as multiplexed illumination [11] and sparse dictionary prior [2] to reconstruct HSI from RGB images which alleviate the need for costly acquisition devices. Despite the advantages of using HSI, it has not been accepted in general vision tasks due to a lack of large scale HSI datasets. Thus, we propose an unsupervised approach for HSI reconstruction from RGB using domain adaptation.

3. Proposed Method

The proposed approach benefits from rich information obtained from different bands in HSI reconstructed from standard RGB images. HSI reconstruction have been achieved via prior-based techniques. However, such an approach is infeasible as the existing aerial image datasets are restricted to infrared, synthetic aperture radar or multispectral images. Therefore, we adopt an approach built upon unsupervised domain adaptation to reconstruct haze-invariant HSI for aerial images.

Given a 3-channel image, $\mathbf{a}_i \in \mathbb{R}^3$ as input, we solve the problem of hallucinating a plausible spectral signature

of the input image. As the problem is under-constrained, we re-formulate the reconstruction task as a spectral reflectance interpolation problem. To achieve this, the individual channels of the input images are spanned to form a 31-channel input. Let an input image $\mathbf{a}_i \in \mathbb{R}^{h \times w \times 3}$. The individual channels in \mathbf{a}_i are spanned to form a spanned image matrix $\mathbf{x}_i \in \mathbb{R}^{h \times w \times 31}$ where h, w denote the image dimensions. The spanned image \mathbf{x}_i is used in turn for generating the HSI.

3.1. Hazy-to-Hyperspectral Reconstruction (H2H)

Given $\{h_i\}_{i=1}^{n_s}$ of HSI sampled from the source domain of natural scenes \mathcal{H} , and $\{x_j\}_{j=1}^{n_t}$ of standard RGB images sampled from the target domain of aerial images \mathcal{X} , the objective of proposed unsupervised domain adaptation framework is to learn how to translate between these domains without paired image examples. The algorithm is based on the assumption that there is a rich underlying HSI dictionary that can be used to reconstruct novel images from the distribution of hyperspectral signatures in natural images. The proposed methodology makes use of several guiding principles such as domain adaptation, task supervision and domain classifier.

Unsupervised Domain Adaptation. First of all, we lack supervision in the form of paired examples. To address the problem, the proposed algorithm exploits supervision at the level of sets. Given a set of HSI and a set of standard RGB images, we train a mapping $G_x(\cdot) : \mathcal{X} \rightarrow \mathcal{H}$ such that the spectral response of the output $\hat{h} = G_x(x), x \in \mathcal{X}$ is indistinguishable from spectral response of $h \in \mathcal{H}$ by an adversary trained to classify \hat{h} apart from h . In addition, we also train a mapping $G_h(\cdot) : \mathcal{H} \rightarrow \mathcal{X}$ which introduces cycle-consistency, thus adding more structure in the proposed algorithm. In context to adversarial losses, the proposed network minimizes the following objective:

$$\mathcal{L}_x(G_x, D_x, x, h) = \mathbb{E}_{x \sim \mathcal{X}}[\log(1 - D_x(G_x(x)))] + \mathbb{E}_{h \sim \mathcal{H}}[\log(D_x(h))] \quad (1)$$

$$\mathcal{L}_h(G_h, D_h, h, x) = \mathbb{E}_{h \sim \mathcal{H}}[\log(1 - D_h(G_h(h)))] + \mathbb{E}_{x \sim \mathcal{X}}[\log(D_h(x))] \quad (2)$$

$$\mathcal{L}_{GAN} = \mathcal{L}_x + \mathcal{L}_h \quad (3)$$

Task Supervision. If we consider 3-channel RGB images in comparison to 31-channel hyperspectral images, hallucinating a plausible spectral response seems daunting. Comparatively, the RGB image has lost 90% of information. Although, the semantics of images and the textural cues provide ample guidance to reconstruct hyperspectral counterparts they are not sufficient to retrieve the actual ground truth. Thus, we aim at learning a plausible HSI image for a given input RGB image by modelling the latent hyperspectral space for image derived from a set of hyperspectral images.

However, due to lack of hyperspectral images for dehazing, we adopt an unpaired approach for hyperspectral reconstruction. For each iteration, an input x is mapped to h by means of a stochastic mapping $x \sim p_{\mathcal{H}, \mathcal{X}}(\hat{h}|x)$. The objective is to learn a plausible spectral response of given RGB image by learning a latent hyperspectral space. Thus, it can be inferred that the $G_x(\cdot)$ can be learned efficiently if $G_x(\cdot)$ is stochastic in nature. This is required to build an output distribution over \hat{h} that reconciles with original distribution $p_{\mathcal{H}}(h)$. However, with limited variation in small HSI datasets, it is problematic to learn an output distribution over \hat{h} that matches its underlying hyperspectral space.

In general, existing approaches address the issue of distributional discrepancy by a form of maximum-mean discrepancy. Rather than relying on a minimax optimization scheme for obtaining aerial HSI, we use our original dehazing task as a self-supervised auxiliary task [39]. Jointly training the HSI reconstruction task along with dehazing task supervision aligns the domains in a subspace specifically for image dehazing. We modify the cycle-consistency loss [49] to incorporate dehazing task supervision. The modified loss encourages $y \approx G_h(G_x(x))$ and $h \approx G_x(G_h(h))$.

$$\mathcal{L}_{cyc} = \|y - G_h(G_x(x))\|_2^2 + \|h - G_x(G_h(h))\|_2^2. \quad (4)$$

where (x, y) refers to hazy and dehazed/clean RGB images respectively, and h refers to HSI.

Domain Classifier. As the primary task of the framework is to produce natural dehazed images, we incorporate a do-

main classifier [10] to learn a mapping $G_x(\cdot)$ which is inherently haze invariant. The motivation is to induce alignment between hazy and clean images by learning a domain classifier as an approximation to the total variation distance. This is achieved by jointly optimizing the algorithm with a domain classifier that discriminates between hazy and clear images during training. By maximizing the loss of domain classifier, the algorithm encourages haze-invariant HSI i.e. \hat{h} in the course of parameter optimization such that the expected target risk $\mathbb{E}_{(x^t) \sim \mathcal{X}}$ is low for loss function $\mathcal{L}_{cls}(\cdot)$,

$$\mathcal{L}_{cls} = \mathbb{E}_{(x^t) \sim \mathcal{X}} \|h - G_x(G_h(h))\|_2^2 \quad (5)$$

where (x, h) refers to hazy RGB image and HSI respectively.

Implementation Details. The generators in G_x , G_h and G_z adopt a U-Net with skip connections while PatchGAN is adopted for the corresponding discriminators. Although the RGB and HSI have considerable difference in the pixel domain, the underlying structure must remain intact. Thus, we use L2 cycle consistency losses and identity losses [50] to further improve the reconstructed HSI. Finally, a ResNet-based architecture is used to extract relevant features from reconstructed HSI, referred to as hyperspectral catalyst, which is further fed into a modified cGAN for efficient dehazing.

3.2. Multi-Cue Image-to-Image Translation (I2I)

To achieve dehazed image outputs in an efficient manner, we use an enhanced cGAN which produces final dehazed output images by utilizing a 12-channel multi-cue color space image as input along with the hyperspectral catalyst.

Multi-cue Color Space. The multi-cue color space is constructed using different color models such as HSV, YCrCb, and LAB concatenated together with RGB which results in a 12-channel input image.

The HSV color model is a cylindrical transformation of RGB Cartesian space in terms of hue, saturation, and luminance value. HSV provides a practical advantage as it separates saturation and luminance, and it has been used extensively in computer vision. Zhang et al. [47] argues that the HSV color space can be used to preserve hue [42] and reduce computational complexity. The hypothesis is based on the fact that HSV specifically models visuals as perceived by human eye which helps in suppressing halo effects.

The YCbCr represents the luminance, blue-difference, and red-difference chroma constituents, respectively. It provides a practical approximation for color processing and perceptual uniformity. The key motivation behind using a YCbCr color space is its robustness towards haze and color

Method	SSIM	PSNR
DCP [16]	0.5430	14.777
CAP [52]	0.6342	10.582
AOD [19]	0.6499	10.227
GFN [36]	0.534	16.787
CycleGAN [50]	0.5950	15.707
GridDehazeNet [23]	0.8915	22.543
SkyGAN	0.8466	24.56

Table 1: Comparative results on HAI dataset

distortion [47, 42]. The use of multi-channel color space enhances the naturalness of the dehazed images [13, 44].

The LAB color space defines a color model that is more perceptually linear than any other color model approximating human perceptual vision. The advantage is especially significant in applications addressing visual degradation as it provides balance corrections to adjust lightness and contrast. Guo et al. [13] and Xie et al. [44] used the Retinex algorithm on the luminance component of the color space to generate the transmission map from a single hazy image.

The proposed method allows the learning algorithm to use different color spaces simultaneously being pivoted on attaining optimal dehazed outputs.

Hyperspectral Catalyst (HSC). It is difficult for the learning algorithm to alleviate problems such as occlusion caused by haze in aerial images. The reconstructed HSI contains spectral information from several wavebands which can mitigate this problem and finally achieve clean aerial images. However, the distribution of values in an HSI vary drastically from that of RGB images. Thus, the reconstructed HSI cannot be directly used for mapping to RGB images.

Therefore, we train a ResNet-based architecture on the reconstructed HSI to obtain relevant feature distribution, which fits the distribution of clean dehazed RGB images. As an analogy to physical-chemical reactions, we refer to the obtained features as a Hyperspectral Catalyst. The HSC is attached with the 12-channel multi-cue input image as additional three channels to form a 15-channel input to cGAN. The ResNet uses the following loss \mathcal{L}_r for parameter optimization:

$$\mathcal{L}_r = \|y - G_r(G_h(G_x(x)))\| \quad (6)$$

where (x, y) refers to hazy RGB image and clean RGB image respectively. G_r denotes the ResNet generator.

4. Dataset and Evaluation

Hazy Aerial Image (HAI) Dataset There has been tremendous development in the area of image restoration

and enhancement, especially with the advent of deep learning. Standardized dataset benchmarks allow fair evaluation of different methodologies under similar evaluation criteria. Several datasets such as Fattal’s dataset [9], FRIDA [41], D-Hazy [1], and the Foggy Cityscapes datasets [37] have been proposed for natural scene image dehazing. Apart from these, the RESIDE Dataset [20] is a benchmark large-scale image dehazing for natural scene images. However, a standard large-scale benchmark for aerial images has been long expected. Recently, Huang et al. [17] proposed the SateHaze1k dataset, which consists of 1200 hazy aerial images.

The HAI Dataset is built upon the Inria Aerial Image Labelling (AIL) Dataset [25]. The AIL Dataset consists of 180 images each in train and test datasets. We crop these images using a sliding overlap with resolution 500×500 , which results in $\sim 65,000$ images in the training set. The HAI dataset is a large-scale dataset for fair evaluation and comparison for aerial image dehazing algorithms. The HAI dataset contains over 65 thousand pairs of synthetic hazy images, generated using 360 high-resolution aerial images. For every image, we create synthetic haze using diamond-square algorithm which generates a plasma fractal or the cloud fractal [31]. The algorithm generates varying degree of vertical and horizontal perturbations. We implement five degrees of simulated haze using a varying grid size. Similarly, we create a test set of 600 images cropped from the AIL test partition.

To corroborate our results, we also evaluate on the recent SateHaze1k dataset [17]. The dataset contains 1200 pairs of hazy images. The images are divided into three levels of fog, namely thin, moderate, and thick fog.

CVPRW NTIRE '18 and NTIRE '20 datasets For HSI reconstruction task, we train our H2H module (Section 3.1) using NITRE 2018’s, ICVL BGU Hyperspectral [2, 3] and NITRE 2020 dataset. ICVL BGU dataset consists of 200 images, from both indoor and outdoor scenes. Along with RGB, they provide 31-channel HSI bands, separated by 10 nm for each image. Similarly, NTIRE 2020 HSI dataset consists of 360 images. For training purposes, the images were augmented by flipping and cropping randomly to generate a total of 6000 images.

4.1. Quantitative Analysis

For evaluating our model quantitatively, we use the test sets from the proposed HAI dataset and SateHaze1k [17] dataset. The results in terms of PSNR and SSIM are compiled in Table 1 and Table 2. The PSNR and SSIM metrics are computed between the de-hazed image and the ground truth clear image.

Performance on HAI Dataset To have a fair comparison of the model on the proposed dataset, we evaluate sev-

Method	Thin Fog		Moderate Fog		Thick Fog	
	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR
Original	0.7241	12.771	0.7399	12.587	0.4215	8.589
DCP [16]	0.7246	13.152	0.5735	9.783	0.585	10.251
SAR-Opt-cGAN [12]	0.8419	20.195	0.7941	21.662	0.7573	19.655
Huang et al. [17] w/o SAR	0.8168	21.474	0.8274	22.095	0.7842	22.121
Huang et al. [17] SAR	0.9061	24.164	0.9264	25.311	0.864	25.073
Dehazenet [5]	0.895	19.753	0.8552	18.125	0.7064	14.332
SkyGAN	0.9248	25.381	0.9035	25.583	0.8925	23.43

Table 2: Comparative results of the proposed method and existing dehazing methods over SateHaze1k [17]

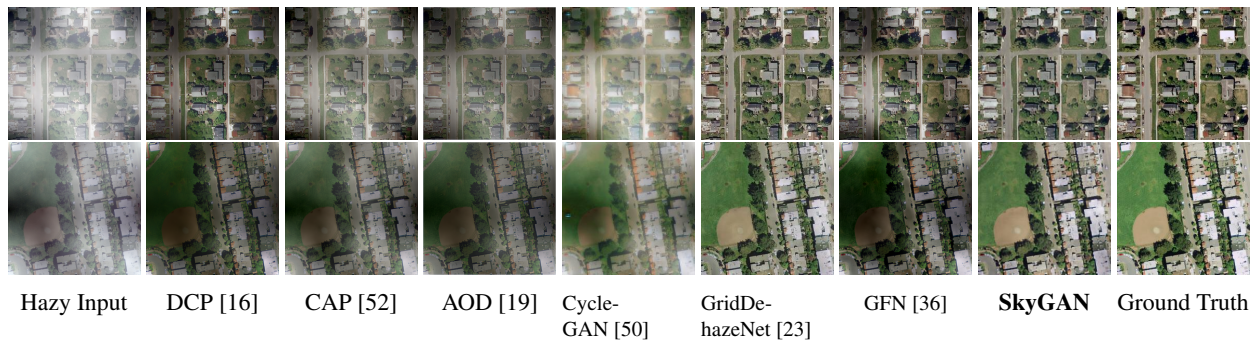


Figure 4: Qualitative comparison on aerial images from HAI Dataset with a city-based scenario

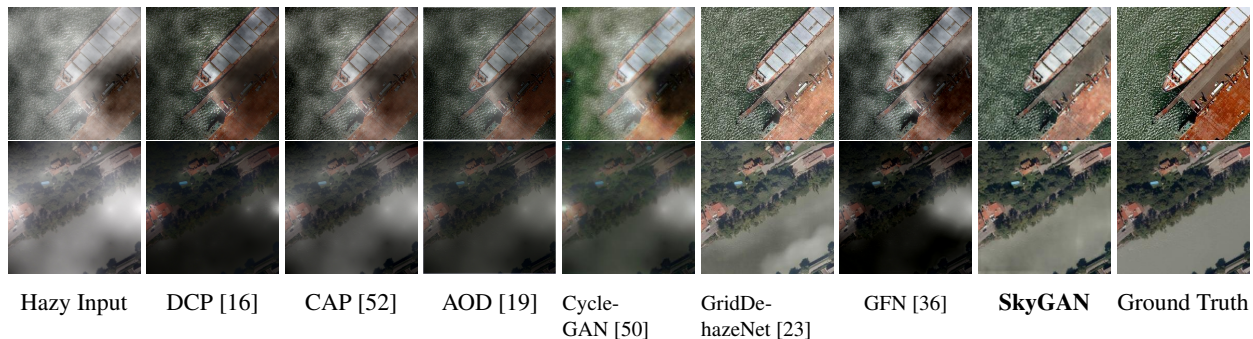


Figure 5: Qualitative comparison on aerial images from HAI Dataset with water bodies and sky regions

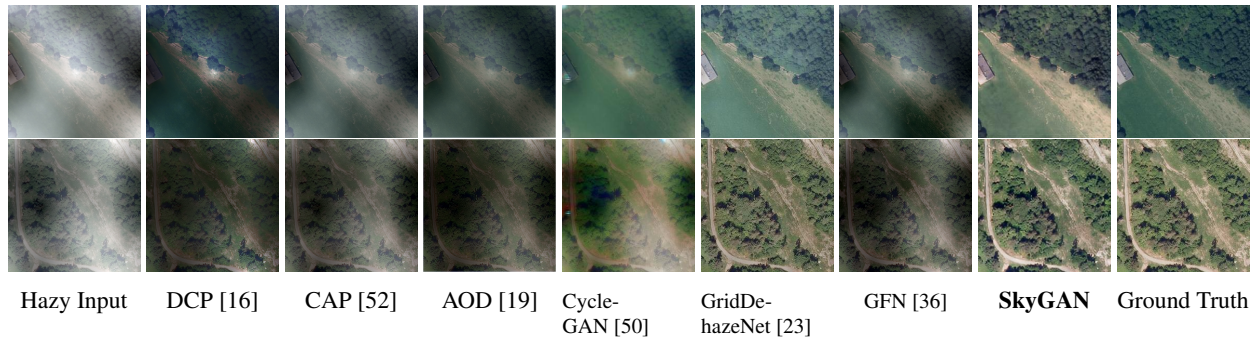


Figure 6: Qualitative comparison on aerial images from HAI Dataset with substantial forest region

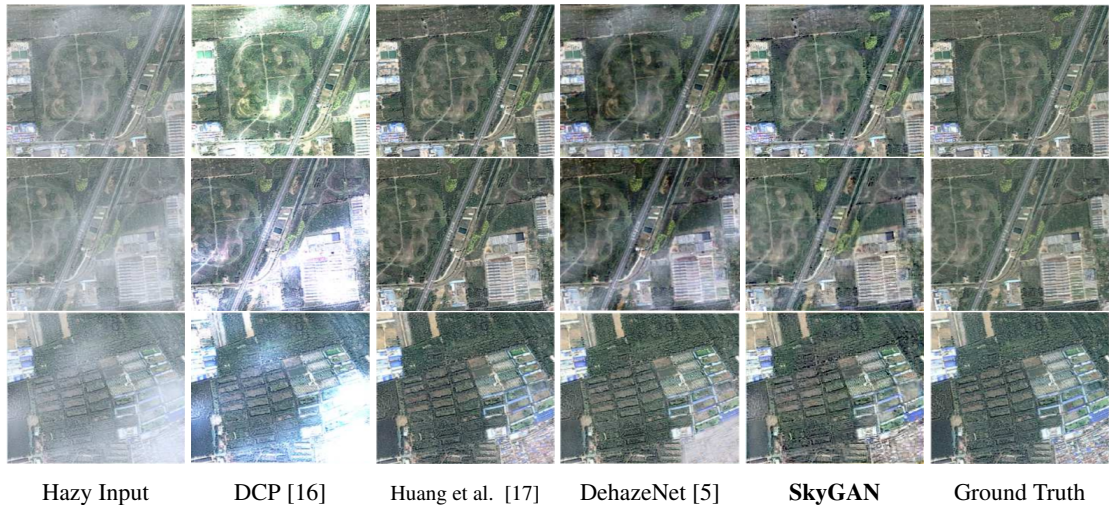


Figure 7: Qualitative comparison of SkyGAN with state-of-the-art techniques over SateHaze1k [17] dataset

eral prior-based and deep learning-based models on HAI dataset. For prior-based techniques, we use the Dark Channel Prior (DCP) [16] and Color Attenuation Prior (CAP) [52]. For deep learning-based techniques, we take into account the state-of-the-art All-in-one dehazing Network (AOD-Net) [19], CycleGAN [50], GFN [36] and GridDehazeNet [23]. The results are presented in Table 1, which indicates the percentage increase SkyGAN could achieve over rest of the techniques. It is quite evident from the metrics that the proposed SkyGAN performs extremely well in comparison to the rest, both in terms of PSNR and SSIM.

Performance on SateHaze1k Dataset. SateHaze1k consists of 45 images from each of the three levels of fog (thin, moderate, and thick). Table 2 shows the quantitative comparison between various techniques and SkyGAN. SkyGAN obtains the highest SSIM in thin and thick fog scenarios. Similarly, our method achieves the best PSNR in thin and moderate fog conditions. Overall, the proposed method outperforms the existing state-of-the-art [17] in 4 out of 6 performance measures (SSIM and PSNR) across the three types of fog conditions.

4.2. Qualitative Analysis

The visual analysis for multiple aerial hazy images was carried out from the HAI and SateHaze1k datasets. From the HAI, a set of six challenging aerial hazy images and the recovered clear images using the proposed SkyGAN and existing methods are compared. To show the robustness of our method, we considered various scenarios from urban regions (in Figure 4), water bodies and sky with large white/blue regions (in Figure 5), and forest regions (in Figure 6). It can be easily assessed that existing methods are

Method	Components			
	Hyperspectral Catalyst	Task Supervision	Domain Classifier	Multi-cue Color Space
Model-1				
Model-2	✓			
Model-3	✓	✓		
Model-4	✓	✓	✓	
SkyGAN	✓	✓	✓	✓

Table 3: Component-wise Ablation Models

able to reduce haze in some regions but fail on certain areas such as grass fields (in Figure 4), red colored ship decks (in Figure 5), and large dense forests (in Figure 6). In contrast, the proposed SkyGAN removes the haze and also restores the color balance in the recovered clear image.

Furthermore, we compare the qualitative results of the proposed and existing methods for three aerial hazy images from SateHaze1k. The images are captured from agricultural fields, roads, residential areas and empty grounds (Figure 7). The proposed method is reducing the haze in the entire image as opposed to the non-uniform improvements obtained by the existing methods. In terms of closeness to the actual ground truth, our results are better than the existing state-of-the-art [17]. The work [17] introduces darker colors in the image as opposed to the actual light colors in the clear images. The higher quantitative performance (PSNR and SSIM) of SkyGAN further proves its robustness to multiple scenarios.

5. Ablation Study

In this subsection, we demonstrate the effectiveness of various components that have successfully contributed to-

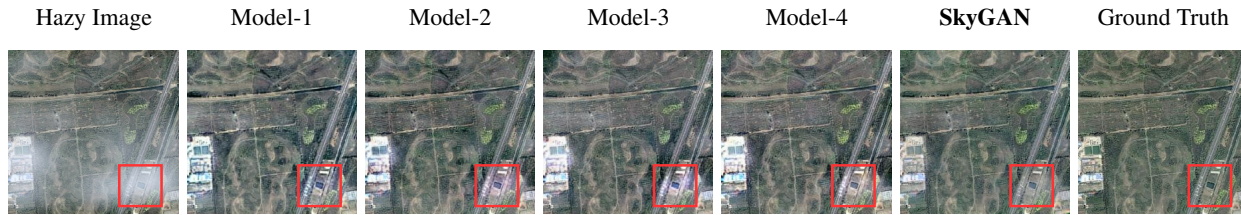


Figure 8: Qualitative comparison for ablation models as described in Table 3. The gradual improvement in visual as highlighted by the red box, clearly indicate the additive enhancement of each component.

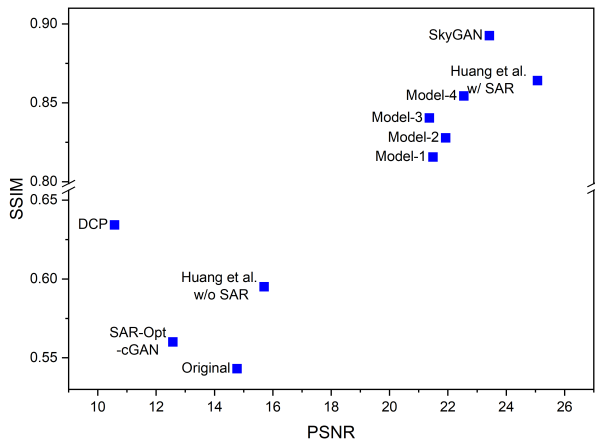


Figure 9: Quantitative comparison for ablation models as described in Table 3 along with different representative state-of-the-art techniques.

wards dehazing the aerial images through a detailed ablation study. This experiment includes 4 progressive models and the SkyGAN as shown in Table 3. We add the various components one-by-one and evaluate their respective performance. We report results on the thick fog level of SateHaze1k dataset for the following ablation models:

Backbone Architecture: The choice of the right image translation model is crucial for the quality of output dehazed images. As cGAN has been widely accepted in literature for a variety of tasks, we choose a modified cGAN as the baseline for our experiments, referred to as Model-1. The successive models differ significantly from Model-1 as it is exclusively based on RGB information. On the other hand, Model 2-4 incorporate HSI information for dehazing. Figure 8 and Figure 9 shows the performance of proposed SkyGAN in comparison to different ablation models along with different representative state-of-the-art techniques for image dehazing respectively.

Hyperspectral Catalyst, Task Supervision, Domain Classifier: To investigate the utility of these auxiliary networks for aerial-image dehazing, we train Model-1 by

adding the components progressively. Figure 9 depicts the performance of various models. Quantitatively, these components leads to increment in performance by 1.5% in terms of SSIM. This can be confirmed by comparing the visual results as shown in Figure 8.

Multi-cue Color Space: In contrast to other components, adding multi-cue color spaces leads to 4.5% improvement in SSIM. This is depicted in Figure 9 as Model 1-4 are clustered in an area, and SkyGAN stands ahead of all these ablation model. Thus multi-cue color space forms a key contributor towards the performance of SkyGAN.

It can be concluded from the comparison, that each component described, makes SkyGAN superior to rest of the models and hence justify their presence.

6. Conclusions

The paper presents a novel GAN based framework, SkyGAN, which utilizes HSI guidance and multi-cue color input for aerial image dehazing. SkyGAN effectively combines adversarial distribution discrepancy alignment and cycle-consistency constraint to reconstruct domain-aware HSI from a set of hyperspectral natural scene images in an unpaired manner. A residual network is further used to alleviate data distribution discrepancy from reconstructed HSI, thereby generating a HSC. The HSC along with a 12-channel multi-cue input is used by the cGAN module (I2I) to perform effective dehazing. The hyperspectral catalyst consists of rich content from several visual bands, which is coupled with the multi-cue color space, leading to highly effective aerial image dehazing. In addition, we also propose a large-scale dataset (HAI) for haze removal in aerial images consisting of realistic, non-homogeneous haze with varying density, along with an extensive evaluation of the dataset on a representative set of existing state-of-the-art techniques. The proposed SkyGAN outperforms the existing methods for aerial image dehazing in terms of PSNR and SSIM on HAI as well as SateHaze1k benchmark.

References

- [1] Cosmin Ancuti, Codruta O Ancuti, and Christophe De Vleeschouwer. D-hazy: A dataset to evaluate quanti-

- tatively dehazing algorithms. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 2226–2230. IEEE, 2016.
- [2] Boaz Arad and Ohad Ben-Shahar. Sparse recovery of hyperspectral signal from natural rgb images. In *European Conference on Computer Vision*, pages 19–34. Springer, 2016.
- [3] 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*, pages 929–938, 2018.
- [4] Dana Berman, Shai Avidan, et al. Non-local image dehazing. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1674–1682, 2016.
- [5] Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao. Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 25(11):5187–5198, 2016.
- [6] Edward A Cloutis. Review article hyperspectral geological remote sensing: evaluation of analytical techniques. *International Journal of Remote Sensing*, 17(12):2215–2242, 1996.
- [7] 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*, pages 0–0, 2019.
- [8] 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*, pages 825–833, 2018.
- [9] Raanan Fattal. Dehazing using color-lines. *ACM transactions on graphics (TOG)*, 34(1):13, 2014.
- [10] Yaroslav Ganin and Victor Lempitsky. Unsupervised domain adaptation by backpropagation. *arXiv preprint arXiv:1409.7495*, 2014.
- [11] Mayank Goel, Eric Whitmire, Alex Mariakakis, T Scott Saponas, Neel Joshi, Dan Morris, Brian Guenter, Marcel Gavrilu, Gaetano Borriello, and Shwetak N Patel. Hypercam: hyperspectral imaging for ubiquitous computing applications. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 145–156, 2015.
- [12] Claas Grohnfeldt, Michael Schmitt, and Xiaoxiang Zhu. A conditional generative adversarial network to fuse sar and multispectral optical data for cloud removal from sentinel-2 images. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 1726–1729. IEEE, 2018.
- [13] Fan Guo, Zixing Cai, Bin Xie, and Jin Tang. Automatic image haze removal based on luminance component. In *2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)*, pages 1–4. IEEE, 2010.
- [14] 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*, pages 0–0, 2019.
- [15] 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*, 90(3):337–352, 2004.
- [16] Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2341–2353, 2010.
- [17] Binghui Huang, Li Zhi, Chao Yang, Fuchun Sun, and Yixu Song. Single satellite optical imagery dehazing using sar image prior based on conditional generative adversarial networks. In *The IEEE Winter Conference on Applications of Computer Vision*, pages 1806–1813, 2020.
- [18] Taeksoo Kim, Moonsu Cha, Hyunsoo Kim, Jung Kwon Lee, and Jiwon Kim. Learning to discover cross-domain relations with generative adversarial networks. *arXiv preprint arXiv:1703.05192*, 2017.
- [19] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. Aod-net: All-in-one dehazing network. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 4770–4778, 2017.
- [20] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing*, 28(1):492–505, 2018.
- [21] Thomas Lillesand, Ralph W Kiefer, and Jonathan Chipman. *Remote sensing and image interpretation*. John Wiley & Sons, 2015.
- [22] Changbing Liu, Jianbo Hu, Yu Lin, Shihong Wu, and Wei Huang. Haze detection, perfection and removal for high spatial resolution satellite imagery. *International Journal of Remote Sensing*, 32(23):8685–8697, 2011.
- [23] Xiaohong Liu, Yongrui Ma, Zhihao Shi, and Jun Chen. Grid-dehazenet: Attention-based multi-scale network for image dehazing. *arXiv preprint arXiv:1908.03245*, 2019.
- [24] Jiao Long, Zhenwei Shi, Wei Tang, and Changshui Zhang. Single remote sensing image dehazing. *IEEE Geoscience and Remote Sensing Letters*, 11(1):59–63, 2013.
- [25] Emmanuel Maggiori, Yuliya Tarabalka, Guillaume Charpiat, and Pierre Alliez. Can semantic labeling methods generalize to any city? the inria aerial image labeling benchmark. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 3226–3229. IEEE, 2017.
- [26] Aliaksei Makarau, Rudolf Richter, Daniel Schläpfer, and Peter Reinartz. Combined haze and cirrus removal for multispectral imagery. *IEEE Geoscience and Remote Sensing Letters*, 13(3):379–383, 2016.
- [27] Murari Mandal, Lav Kush Kumar, and Santosh Kumar Vipparthi. Mor-uav: A benchmark dataset and baselines for moving object recognition in uav videos. *arXiv preprint arXiv:2008.01699*, 2020.
- [28] Murari Mandal, Manal Shah, Prashant Meena, Sanhita Devi, and Santosh Kumar Vipparthi. Avdnet: A small-sized vehicle detection network for aerial visual data. *IEEE Geoscience and Remote Sensing Letters*, 2019.

- [29] 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)*, pages 3098–3102. IEEE, 2019.
- [30] Aditya Mehta, Harsh Sinha, Pratik Narang, and Murari Mandal. Hidegan: A hyperspectral-guided image dehazing gan. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 212–213, 2020.
- [31] Claudio Michaelis, Benjamin Mitzkus, Robert Geirhos, Evgenia Rusak, Oliver Bringmann, Alexander S. Ecker, Matthias Bethge, and Wieland Brendel. Benchmarking robustness in object detection: Autonomous driving when winter is coming. *arXiv preprint arXiv:1907.07484*, 2019.
- [32] Lichao Mou, Yuansheng Hua, Pu Jin, and Xiao Xiang Zhu. Era: A dataset and deep learning benchmark for event recognition in aerial videos. *arXiv preprint arXiv:2001.11394*, 2020.
- [33] 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*, 3:251–307, 1999.
- [34] 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*, pages 8160–8168, 2019.
- [35] 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*, pages 8160–8168, 2019.
- [36] Wenqi Ren, Lin Ma, Jiawei Zhang, Jinshan Pan, Xiaochun Cao, Wei Liu, and Ming-Hsuan Yang. Gated fusion network for single image dehazing. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2018.
- [37] Christos Sakaridis, Dengxin Dai, and Luc Van Gool. Semantic foggy scene understanding with synthetic data. *International Journal of Computer Vision*, 126(9):973–992, 2018.
- [38] Huanfeng Shen, Chi Zhang, Huifang Li, Quan Yuan, and Liangpei Zhang. A spatial-spectral adaptive haze removal method for visible remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 2020.
- [39] Yu Sun, Eric Tzeng, Trevor Darrell, and Alexei A Efros. Unsupervised domain adaptation through self-supervision. *arXiv preprint arXiv:1909.11825*, 2019.
- [40] Robby T Tan. Visibility in bad weather from a single image. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2008.
- [41] Jean-Philippe Tarel, Nicholas Hautiere, Laurent Caraffa, Aurélien Cord, Houssam Halmaoui, and Dominique Gruyer. Vision enhancement in homogeneous and heterogeneous fog. *IEEE Intelligent Transportation Systems Magazine*, 4(2):6–20, 2012.
- [42] Yi Wan and Qiqiang Chen. Joint image dehazing and contrast enhancement using the hsv color space. In *2015 Visual Communications and Image Processing (VCIP)*, pages 1–4. IEEE, 2015.
- [43] Gui-Song Xia, Xiang Bai, Jian Ding, Zhen Zhu, Serge Belongie, Jiebo Luo, Mihai Datcu, Marcello Pelillo, and Liangpei Zhang. Dota: A large-scale dataset for object detection in aerial images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3974–3983, 2018.
- [44] Bin Xie, Fan Guo, and Zixing Cai. Improved single image dehazing using dark channel prior and multi-scale retinex. In *2010 International Conference on Intelligent System Design and Engineering Application*, volume 1, pages 848–851. IEEE, 2010.
- [45] Zili Yi, Hao Zhang, Ping Tan, and Minglun Gong. Dualgan: Unsupervised dual learning for image-to-image translation. In *Proceedings of the IEEE international conference on computer vision*, pages 2849–2857, 2017.
- [46] He Zhang and Vishal M Patel. Densely connected pyramid dehazing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3194–3203, 2018.
- [47] Tianyu Zhang, Hai-Miao Hu, and Bo Li. A naturalness preserved fast dehazing algorithm using hsv color space. *IEEE Access*, 6:10644–10649, 2018.
- [48] Ying Zhang and Bert Guindon. Quantitative assessment of a haze suppression methodology for satellite imagery: Effect on land cover classification performance. *IEEE Transactions on Geoscience and Remote Sensing*, 41(5):1082–1089, 2003.
- [49] Tinghui Zhou, Philipp Krahenbuhl, Mathieu Aubry, Qixing Huang, and Alexei A Efros. Learning dense correspondence via 3d-guided cycle consistency. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 117–126, 2016.
- [50] 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*, pages 2223–2232, 2017.
- [51] Lin Zhu, Yushi Chen, Pedram Ghamisi, and Jón Atli Benediktsson. Generative adversarial networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 56(9):5046–5063, 2018.
- [52] Qingsong Zhu, Jiaming Mai, and Ling Shao. A fast single image haze removal algorithm using color attenuation prior. *IEEE transactions on image processing*, 24(11):3522–3533, 2015.