

Unsupervised Martian Dust Storm Removal via Disentangled Representation Learning

Dong Zhao^{1,2}, Jia Li^{1,3,4}, Hongyu Li¹, and Long Xu^{2,3}

¹ State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University, Beijing 100191, China

jiali@buaa.edu.cn

² State Key Laboratory of Space Weather, National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China ³ Peng Cheng Laboratory, Shenzhen 518000, China

Received 2022 May 21; revised 2022 June 30; accepted 2022 July 11; published 2022 August 30

Abstract

Mars exploration has become a hot spot in recent years and is still advancing rapidly. However, Mars has massive dust storms that may cover many areas of the planet and last for weeks or even months. The local/global dust storms are so influential that they can significantly reduce visibility, and thereby the images captured by the cameras on the Mars rover are degraded severely. This work presents an unsupervised Martian dust storm removal network via disentangled representation learning (DRL). The core idea of the DRL framework is to use the content encoder and dust storm encoder to disentangle the degraded images into content features (on domain-invariant space) and dust storm features (on domain-specific space). The dust storm features carry the full dust storm-relevant prior knowledge from the dust storm images. The "cleaned" content features can be effectively decoded to generate more natural, faithful, clear images. The primary advantages of this framework are twofold. First, it is among the first to perform unsupervised training in Martian dust storm removal with a single image, avoiding the synthetic data requirements. Second, the model can implicitly learn the dust storm-relevant prior knowledge from the real-world dust storm data sets, avoiding the design of the complicated handcrafted priors. Extensive experiments demonstrate the DRL framework's effectiveness and show the promising performance of our network for Martian dust storm removal.

Key words: planets and satellites: atmospheres - planets and satellites: general - planets and satellites: surfaces

1. Introduction

The Zhurong rover, China's first Mars rover, went into hibernation mode on 2022 May 18, to cope with the reducing solar power generation capacity caused by a dust storm which was observed by the medium resolution images obtained from the orbiter of the Tianwen-1 probe on 2022 March 16 and April 30. That is one case of the Martian dust storms posing a threat to the scientific instruments (Alexandra Witze 2018). Unfortunately, Martian dust storms are common, which may last for weeks and cover many areas of the planet (Clarke 2018). These dust storms are so influential that they can change the climate of Mars (Leovy 2001; Liu et al. 2022) and degrade signal/image qualities taken by cameras on the Mars rover. For example, as shown in Figure 1, the images captured under the dust storm weather suffer from contrast degradation, color attenuation, and poor visibility. Therefore, removing the dust storm from a single image becomes crucial for studying the Martian atmosphere (Banfield et al. 2020) and Martian geology (Chaffin et al. 2021).

To this end, we have recalled many related studies on image enhancement and restoration algorithms. Among them, we focus on the haze removal task because it is very similar to dust storm removal, where the main difference between them is the suspended particles in the atmosphere. These image restoration approaches are categorized into five groups as follows.

Conventional image enhancement techniques, such as Auto Levels (AL), Retinex model (Jobson et al. 1997) or contrast limited adaptive histogram equalization (CLAHE) (Reza 2004) can improve the visual quality of the Martian dust storm images. These methods have a lower computational cost, yet some of their common drawbacks are that some parameters should be manually configured. Thus, they cannot automatically adapt to different images captured under different sizes of dust storms.

Physical prior-based method is another kind of approach that might work. For example, inspired by the success of the Earth's atmospheric scattering model (Narasimhan 2000) that has been leveraged in natural image dehazing by introducing specific priors (e.g., dark channel prior (DCP) (He et al. 2011), colorlines prior (Fattal 2014), color attenuation prior (Zhu et al. 2015), and non-local prior (NLP) (Berman et al. 2018), Li et al. (2018) directly utilized the DCP for Martian dust storm removal. However, the sizes, types, and concentrations in the space of the aerosol particles on Mars are very different from the ones on Earth; therefore, the atmospheric scattering models

⁴ Corresponding author.



Figure 1. The motivation of the DRL for unsupervised Martian dust storm removal.

of the two planets are different (Egan & Foreman 1971). For example, typically, Rayleigh scattering occurs on Earth, while Mie scattering on Mars (Collienne et al. 2013). That means the existing priors proposed, especially for nature image dehazing on Earth, are unable to fit all conditions on Mars and may produce unwanted artifacts because the priors are invalid.

In contrast, supervised learning-based approaches can learn the latent prior knowledge from masses of paired contaminated images and their clear counterparts and produce visually appealing results. For example, in the multi-scale CNNs (MSCNN) (Ren et al. 2020) and enhanced pix2pix dehazing networks (EPDN) (Qu et al. 2019), the hazy samples are often synthesized by applying the atmospheric scattering model. Despite their effectiveness on the open-access data sets, supervised learning-based methods have suffered from the domain shift issues of the synthetic data sets. Training on such data sets, these methods are probably to overfit. Therefore, they are less able to generalize well to real-world images.

To improve the performances in the real-world conditions, some semi-supervised learning-based models, such as semisupervised dehazing network (SSDN) (Li et al. 2019) and domain adaptation dehazing network (DADN) (Shao et al. 2020), have been exploited to use both synthetic and real-world data sets. However, as to our task, such paired dust storm images and clear images are impractical to collect under realworld conditions on Mars.

Unlike the above learning-based models, unsupervised learning-based models are solely trained on the real-world data set, avoiding labor-intensive collecting data and dealing with the domain shift problem. However, due to the lack of prior knowledge, the existing unsupervised models inevitably exploit the properties of clear images via predetermined prior losses. For example, the performances of deep dark channel prior (DDCP) (Golts et al. 2019) and zero-shot image dehazing (ZID) (Li et al. 2020) are very dependent on the dark channel prior (He et al. 2011) loss. Like the physical prior-based methods, using the handcrafted priors as the objections may lead to the networks being less robust to the conditions that the priors are invalid.

By discussing the approaches above, we found that the unsupervised learning-based method is more fit for the Martian dust storm removal task for two reasons. First, this method does not need paired clear and dust storm images. Second, it can implicitly learn the dust storm-relevant prior knowledge from real-world dust storm data sets. To this end, we develop an unsupervised Martian dust storm removal framework that learns the dust storm-relevant prior knowledge via disentangled representation learning (DRL). The motivation of the DRL can be succinctly illustrated in Figure 1, where the "disentanglement" module aims to encode the input dust storm image into intermediate representations, i.e., content features (domaininvariant cues) and dust storm features (domain-specific cues). Because the dust storm features are "taken away" from the original image, the remained content features can be effectively decoded into a clear image. The dust storm features carry the full dust storm-relevant prior knowledge of the original image, and they can be fused to another clear image's content features to generate a new dust storm image. Concretely, the proposed



Figure 2. The latent space assumption in the DRL. The clear and dust storm images belong in two different domains C and D, respectively. They can be mapped to content features e^{c}/c^{d} in domain-invariant space, while dust storm features d in domain-specific space can be disentangled from domain \mathcal{D} .

framework consists of two content encoders to extract content features from unpaired clear and dust storm images, one dust storm encoder to disentangle dust storm features, two generators to reconstruct clear and dust storm images, and two discriminators to align the generated images into their corresponding domains. We use the adversarial loss during the training to align the content features to the ones of clear image. In addition, the cross-cycle consistency loss is used to guarantee the content consistency between the original and the reconstructed images. Finally, we also use a latent reconstruction loss to encourage bidirectional mapping of the intermediate representations. Extensive experiments show the promising performance of our Martian dust storm removal framework using DRL.

This paper is structured as follows. In Section 2, the details of our unsupervised Martian dust storm removal network using disentangled representation learning are introduced. We then introduce our data sets in Section 3. We further report the experimental results and the ablation studies in Section 4. In Section 5, we summarize our approach.

2. Method

The DRL can be used to model one or more factors of domain variations while the other factors remain relatively invariant. Specifically, in the Martian dust storm removal task, there are two domains: the clear image domain C and dust storm image domain \mathcal{D} . The goal of the DRL for unsupervised dust storm removal is to learn the domain-specific factors (i.e., dust storm features d) from domain \mathcal{D} and the domain-invariant factors (content features c^c and c^d) from domain C and domain \mathcal{D} , as depicted in Figure 2. To do so, the DRL model in our work consists of one dust storm encoder E_{ds} to extract the dust storm features; two content encoders E_{ct}^c and E_{ct}^d to extract the content features c^c and c^d from clear and dust storm images, respectively; and two image generators G^{c} and G^{d} to map the

inputs onto clear and dust storm image domains, respectively. Details on the network architectures and training objectives are introduced in the following subsections.

2.1. Network Architecture

2.1.1. Overview

The overall disentangled representation learning-based framework for the Martian dust storm removal consists of three parts: forward translations, backward reconstructions, and self-reconstructions, as shown in Figure 3. The notations in this figure are summarized in Table 1.

Forward Translations: The forward translation contains two branches: one for the forward dust storm image translation and the other for the forward clear image translation.

For forward dust storm image translation, the content encoder E_{ct}^c maps clear image I^c onto a shared, domaininvariant content space, and the dust storm encoder E_{ds} maps dust storm image I^d onto a domain-specific space—the named dust storm space in this work. Then, the dust storm generator G^d takes clear content features $c^c = E_{ct}^c(I^c)$ conditioned on dust storm features $d = E_{ds}(I^d)$ to generate dust storm image I_{fwd}^d :

$$I_{\rm fwd}^d = G^d(c^c, d) = G^d(E_{\rm ct}^c(I^c), E_{\rm ds}(I^d)).$$
(1)

For forward clear image translation, the content encoder E_{ct}^d maps dust storm images I^d onto the shared content space. The dust storm removal generator G^c takes dust storm content features $\bar{d} = E_{ct}^d(I^d)$ to generate clear image I_{fwd}^c :

$$I_{\text{fwd}}^c = G^c(\boldsymbol{c}^d) = G^c(E_{\text{ct}}^d(I^d)).$$
⁽²⁾

Backward Reconstructions: The stage of backward reconstructions is similar to the forward translations, where the input images are I_{fwd}^d and I_{fwd}^c .



(a) Forward Translations and Backward Reconstructions



(b) Self-Reconstructions for clear image (left) and dust storm image (right)

Figure 3. Overview of the unsupervised Martian dust storm removal network using disentangled representation learning.

For backward dust storm image reconstruction, we have:

For backward clear image reconstruction, we have:

$$I_{\text{bwd}}^c = G^c(\bar{c}^d) = G^c(E_{\text{ct}}^d(I_{\text{fwd}}^d)).$$
(4)

 $I_{\text{bwd}}^d = G^d(\bar{\boldsymbol{c}}^c, \bar{\boldsymbol{d}}) = G^d(E_{\text{ct}}^c(I_{\text{fwd}}^c), E_{\text{ds}}(I_{\text{fwd}}^d)),$ (3)

where I_{fwd}^d and I_{fwd}^c are the clear and dust storm images defined in (1) and (2), respectively; the \bar{c}^c and \bar{d} are the content features of image I_{fwd}^c and the dust storm features of image I_{fwd}^d , respectively. where the \bar{c}^d represents the content feature of image I_{fwd}^d .

Self-Reconstructions: To facilitate the training process, we apply self-reconstructions as illustrated in Figure 3(b). With encoded content and dust storm features, the generators G^c and G^d should decode them back to original inputs I^c and I^d ,

Notation	Description	Notation	Description		
$E_{\rm ct}^c/E_{\rm ct}^d$	content encoder for the clear/dust storm image	D^c/D^d	discriminator for clear/dust storm domain		
Eds	dust storm encoder	I^c/I^d	input clear/dust storm image		
G^c/G^d	clear/dust storm image generator	$I^d_{ m fwd}/I^d_{ m bwd}$	dust storm image translated from the clear image $I^c/I^c_{\rm fwd}$		
c^{c}/c^{d}	content feature of the clear/dust storm image I^c/I^d				
$ar{m{c}}^c/ar{m{c}}^d$	content feature of the clear/dust storm image $I_{\rm fwd}^c/I_{\rm bwd}^d$	$I_{ m fwd}^c/I_{ m bwd}^c$	clear image restored form the dust storm image $I^d/I^d_{\rm fwd}$		
d	dust storm feature of the dust storm image I^d				
ā	\bar{d} dust storm feature of the dust storm image $I_{\rm fwd}^d$		self-reconstructed clear/dust storm image		

 Table 1

 A Summary of the Notations used in Our Unsupervised Martian Dust Storm Removal Framework

respectively. They can be formulated by:

$$I_{\text{self}}^c = G^c(E_{\text{ct}}^c(I^c));$$

$$I_{\text{self}}^d = G^d(E_{\text{ct}}^d(I^d), E_{\text{ds}}(I^d)).$$
(5)

2.1.2. Details

We adapt the above architecture for the unsupervised framework with the following configurations.

Identity Shortcut Connection and Long Skip Connection: The details of each layer in the encoder and generator are illustrated in Figure 4. Each basic layer in the encoders/ generators consists of a convolutional/deconvolution for down/up-sampling followed by a ResNet (He et al. 2016) block to avoid the notorious exploding or vanishing gradients that may occur during the training. One ResNet block is realized by inserting shortcut connections into conventional convolutional layers, as shown in Figure 4. Note that the shortcut connections introduce neither extra parameter nor computation complexity.

We also equip long skip connections between encoders and generators, referred to as U-shape networks (U-Net) (Ronneberger et al. 2015), as shown in Figure 4. The long skip connections allow the network to propagate structure (at lowlevel) and semantic (at high-level) information of inputs to deeper layers, preserving the spatial information lost during downsampling in the encoders.

Sparse Switchable Normalization and Adaptive Instance Normalization: We use the sparse switchable normalization (SSN) (Shao et al. 2019) in the basic layer instead of the batch normalization (BN) (Ioffe & Szegedy 2015) and instance normalization (IN) (Ulyanov et al. 2017) that are widely used in natural image generation networks. Our earlier explorations found that using BN and IN alone would lead to "spot" artifacts in the generated images, as illustrated in Figure 8. One reason might stem from challenges inherent in the Martian dust storm removal networks where different layers should behave differently. SSN can address this issue by adaptively selecting the optimal normalizers among BN, IN, and layer normalization (LN) (Ba et al. 2016) at different convolutional layers.

In our model, the dust storm image generator G^d fuses the content features and dust storm features via adaptive instance normalization (AdaIN) (Huang & Belongie 2017). The clear content features c^c and dust storm features d are fed into the AdaIN layer. The AdaIN layer aims to align the content features' mean and variance to the ones of dust storm features:

$$\boldsymbol{f} = \delta(\boldsymbol{d}) \cdot \frac{\boldsymbol{c}^c - \mu(\boldsymbol{c}^c)}{\delta(\boldsymbol{c}^c)} + \mu(\boldsymbol{d}), \tag{6}$$

where $\mu(\cdot)$ and $\delta(\cdot)$ are the mean and standard deviation operations.

2.2. Training Objective

The overall loss function of the proposed network is:

$$\mathcal{L} = \lambda_{adv} \mathcal{L}_{adv} + \lambda_{ccc} \mathcal{L}_{ccc} + \lambda_{src} \mathcal{L}_{src} + \lambda_{lrc} \mathcal{L}_{lrc}, \qquad (7)$$

where \mathcal{L}_{adv} , \mathcal{L}_{ccc} , \mathcal{L}_{src} and \mathcal{L}_{lrc} are the adversarial loss, crosscycle consistency loss, self-reconstruction loss and latent reconstruction loss, respectively. λ_{adv} , λ_{ccc} , λ_{src} and λ_{lrc} are the corresponding hyper-parameters to control the importance of each term.

Adversarial Loss: Here, we apply the adversarial loss (Goodfellow et al. 2014) between the generated images and corresponding target domains to force them to have as similar data distributions as possible. For the dust storm image domain, we define the adversarial loss as:

$$\mathcal{L}_{adv}(E_{ct}^c, E_{ds}, G^d) = \mathbb{E}_{I^d}[\log \mathsf{D}^d(I^d)] + \mathbb{E}_{I^c, I^d}[\log(1 - \mathsf{D}^d(G^d(E_{ct}^c(I^c), E_{ds}(I^d))))],$$
(8)

where D^d is the discriminator for the dust storm image domain. \mathbb{E} represents the mean operation of training samples in a batch. Our discriminator module is "PatchGAN" (Isola et al. 2017) classifiers.



Figure 4. Details of the encoders $(E_{ct}^c, E_{ds}, \text{and } E_{ct}^d)$ and generators $(G^d \text{ and } G^c)$ (taking the *forward translation* as an example). \bigcirc : elemental-wise subtraction; \oslash : elemental-wise division; \otimes : elemental-wise multiplication; \oplus : elemental-wise addition.



Figure 5. Examples of the available samples in the data sets.

Similarly, for the clear image domain, the adversarial loss is defined as:

$$\mathcal{L}_{adv}(E_{ct}^d, G^c) = \mathbb{E}_{I^c}[\log \mathsf{D}^c(I^c)] + \mathbb{E}_{I^d}[\log(1 - \mathsf{D}^c(G^c(E_{ct}^d(I^d))))], \qquad (9)$$

where D^c is the discriminator for clear image domain.

Cross-Cycle Consistency Loss: We use the cross-cycle consistency loss to guarantee that the clear images I_{fwd}^c can be reconstructed to the dust storm domain, and the dust storm image I_{fwd}^d can be reconstructed back to the clear domain. Specifically, the cross-cycle consistency loss is defined as:

$$\mathcal{L}_{ccc}(E_{c}^{c}, E_{ds}, E_{c}^{d}, G^{d}, G^{c}) = \mathbb{E}_{I^{d}, I^{c}} \left[\| G^{d}(E_{ct}^{c}(I_{fwd}^{c}), E_{ds}(I_{fwd}^{d})) - I^{d} \|_{2} + \| G^{c}(E_{ct}^{d}(I_{fwd}^{d})) - I^{c} \|_{2} \right],$$
(10)

where I_{fwd}^d and I_{fwd}^c are defined as (1) and (2), respectively.

Self-Reconstruction Loss: With encoded content and dust storm features, the generators G^d and G^c should decode them back to the corresponding domains. Thus, to facilitate the training process, here we use a self-reconstruction loss as illustrated in Figure 3(b):

$$\mathcal{L}_{\rm src}(E_{\rm ct}^{c}, E_{\rm ds}, E_{\rm ct}^{d}, G^{d}, G^{c}) \\= \mathbb{E}_{I^{d}, I^{c}} [\|G^{d}(E_{\rm ct}^{d}(I^{d}), E_{\rm ds}(I^{d})) \\\times -I^{d}\|_{2} + \|G^{c}(E_{\rm ct}^{c}(I^{c})) - I^{c}\|_{2}].$$
(11)

Latent Reconstruction Loss: To encourage the invariant representation learning between the clear and the dust storm space, we apply a latent reconstruction (including content and dust storm features reconstruction) loss similar to Lin et al. (2018) and Zhu et al. (2017) as follows:

$$\mathcal{L}_{\rm Irc}(E_{\rm ct}^{c}, E_{\rm ds}, E_{\rm ct}^{d}, G^{d}, G^{c}) = \mathbb{E}_{I^{c}, I^{d}, I_{\rm fwd}^{c}} [\omega^{c^{c}} \cdot \|E_{\rm ct}^{c}(I^{c}) - E_{\rm ct}^{d}(I_{\rm fwd}^{d})\|_{l} + \omega^{c^{d}} \cdot \|E_{\rm ct}^{d}(I^{d}) - E_{\rm ct}^{c}(I_{\rm fwd}^{c})\|_{l} + \omega^{d} \cdot \|E_{\rm ds}(I^{d}) - E_{\rm ds}(I_{\rm fwd}^{d})\|_{l}],$$
(12)

where ω^{c^c} , ω^{c^d} and ω^d are weights of the features for corresponding reconstructions, respectively. Note that the content features are more easily learned than the dust storm features. It is because the formers can be aligned by the long skip connections, delivering multi-scales of content information from encoders to the corresponding generators. In contrast, the dust storm features are only learned at the single high-level layer. To address this discrepancy, the ω^d is set larger than ω^{c^c} and ω^{c^d} .

3. Dataset

The Mars32k Dataset (Dominik Schmidt 2018) collects 32,368 560×500 samples captured by the Curiosity rover between 2012 and 2018 on Mars. All samples are provided by NASA/JPL-Caltech. This data set involves a variety of geographic features of Mars, including rocky terrain, sand dunes, and mountains. However, we found that the raw Mars32k data set cannot be directly served as the training data for our task accounts for two aspects. First, the raw data set is not curated, containing many anomalous samples. Second, it is difficult for some scenes to determine whether they have a dust storm or not. To build the available data sets, we first artificially eliminate the anomalous samples and select the available samples that are confirmed to have dust storms or are clear. After going through the above procedures, we obtain 1925 dust storm samples and 1432 clear samples, as shown in Figure 5. Then, we randomly selected 65 dust storm samples as the testing data sets and remained samples were used for training data augmentation. We augment the samples by flipping, randomly cropping, and randomly rotating. Finally, the data set contains 13,016 dust storms and 8985 clear unpaired samples for model training. Considering the limits of memory and fast training, we resize the input samples to 256×256 . While during the testing, the inputs are resized to 512×512 .

4. Experiments

4.1. Metrics

Because it is impractical to acquire Martian dust storm images and their clear counterparts, we use the following three no-reference image quality metrics (NRIQM) in our evaluation experiments.

Geometric Mean of the Ratios of Visibility Level (GMRVL) (Hautiere et al. 2008). The GMRVL is a function of the ratio between the visibility levels of restored and original images. It is used to measure the quality of the contrast restoration.

Blind/Referenceless Image Spatial Quality Evaluator (BRIS-QUE) (Mittal et al. 2012). The BRISQUE is a natural scene statistic-based NRIQM calculated in the spatial domain. It can quantify the quality of distortion and naturalness using locally normalized luminance. We use this metric to evaluate whether the dust storm images have been satisfactorily restored with good visual perception.

Perception-based Image Quality Evaluator (PIQE) (Venkatanath et al. 2015). The PIQE extracts local features for measuring patch-wise restoration quality. It is an opinionunaware methodology and thus does not need any training data. It can measure whether the images are satisfactorily restored: insufficient restored or over-restored results would lead to a higher PIQE score.

4.2. Implemental Details

We develop our model on PyTorch deep learning package. We adopt Adam (Kingma & Ba 2015) as the optimization algorithm during the training, with a batch size of 8. In the experiments, our model is trained for 20 epochs, taking about 36 hr on two NVIDIA GeForce GTX 3090 Ti GPU. We set the initial learning rate is 0.0002 for all of the encoders, generators and discriminators, and the learning rate decay is $\gamma = 0.5$ for every 10 epochs. The parameters of the hybrid loss functions are set as: $(\lambda_{adv}, \lambda_{ccc}, \lambda_{src}, \lambda_{lrc}) = (0.2, 3, 1, 0.5)$. Weights of the dust storm features are set as: $(\omega^{e^c}, \omega^{e^d}, \omega^d) = (1, 1, 10)$. The model will be available as open source at https://github. com/phoenixtreesky7/DRL_UMDSR.

4.3. Comparisons with Other Image Enhancement/ Dehazing Methods

We compare our model with the following methods: (1) Conventional Image Enhancement methods (AL and CLAHE Reza 2004); (2) Physical Prior-based methods (DCP He et al. 2011 and NLP Berman et al. 2018); (3) Supervised Learningbased method (MSCNN Ren et al. 2020 and EPDN Qu et al. 2019); (4) Semi-supervised Learning-based methods (SSDN Li et al. 2019 and DADN Shao et al. 2020); (5) Unsupervised Learning-based methods (ZID Li et al. 2020). Qualitative and quantitative results are illustrated in Figure 6 and Table 2, respectively.

Qualitative Comparisons: All algorithms can be observed to restore the dust storm images with good visual perception to some extent. However, the visual results of AL and DADN (Shao et al. 2020) contribute false color, more like the photos taken on Earth. The CLAHE (Reza 2004) and SSDN (Li et al.

2019) seem to fail to remove the dust storm effectively, producing poor visual results in distant scenes. Although the physical prior-based methods DCP (He et al. 2011) and NLP (Berman et al. 2018) can effectively remove dust storms from the scenes, they tend to yield over-enhanced visual artifacts, especially in the sky regions. These results also demonstrate that these priors are invalid for Martian dust storm images. The results of the MSCNN (Ren et al. 2020) and EPDN (Qu et al. 2019) maintain plausible visual details, achieving better performance than the non-learning methods. However, some visual results are not satisfactory because some scenes are oversaturation, as illustrated in the first, the fifth, and the last images in Figure 6 (g). The restored images of ZID (Li et al. 2020) have severe color distortion in the sky regions. By contrast, our model achieves remarkable performance. Note that our visual results may not be as colorful as the ones of AL, DCP (He et al. 2011), MSCNN (Ren et al. 2020), and DADN (Shao et al. 2020). However, our DRL framework extracts the dust storm factors from the dust storm images, and the remaining content features are aligned to the ones of real-world Martian clear images. It enables the restored image to display consistent and faithful colors. Moreover, our model generates high-quality restored results with fewer artifacts than all the compared methods, as illustrated in Figure 7 where the red blocks are the noticeable artifacts masks detected by PIQE (Venkatanath et al. 2015) metric.

Quantitative Comparisons: We further use the GMRVL (Hautiere et al. 2008), BRISQUE (Mittal et al. 2012) and PIQE (Venkatanath et al. 2015) to quantitatively evaluate the effectiveness of our model. The results are listed in Table 2. It can be found that the conventional image enhancement models can generate high contrast results with higher GMRVL than learning-based methods. However, the other two metrics of them are worse, indicating that the results are unnatural and also demonstrating that these methods are unable to cope with complex dust storm conditions. The physical prior-based methods can generally obtain better GMRVL scores than other methods. However, they are more likely to lead to unnatural results with more artifacts as they get high BRISQUE and PIQE. The reasons are that these priors are invalid on Mars. The supervised and semi-supervised methods successfully refrain from over-saturation, resulting in smaller BRISQUE and PIQE than the non-learning methods. However, despite their superiorities, their BRISQUE and PIQW scores are still higher (worse) than ours. Thanks to the DRL, our model achieved better visual quality and obtained the lowest scores of the BRISQUE and PIQE. As Table 2 demonstrates, our model surpasses the second-best methods with gains of 1.123 on BRISQUE and 5.517 on PIQE. Note that our GMRVL is lower than some of the compared methods. The reason is that these methods tend to outcome over-enhanced results, which may increase the GMRVL to a certain extent.



Figure 6. Visual comparisons with other methods, including (b) AL, (c) CLAHE (Reza 2004), (d) DCP (He et al. 2011), (e) NLP (Berman et al. 2018), (f) MSCNN (Ren et al. 2020), (g) EPDN (Qu et al. 2019), (h) SSDN (Li et al. 2019), (i) DADN (Shao et al. 2020) and (j) ZID (Li et al. 2020). (a) The dust storm images. (k) The results of our method.

Zhao et al.

Figure 7. The noticeable artifacts masks detected by PIQE (Venkatanath et al. 2015) metric for (b) AL, (c) CLAHE (Reza 2004), (d) DCP (He et al. 2011), (e) NLP (Berman et al. 2018), (f) MSCNN (Ren et al. 2020), (g) EPDN (Qu et al. 2019), (h) SSDN (Li et al. 2019), (i) DADN (Shao et al. 2020), (j) ZID (Li et al. 2020) and (k) ours results. (a) The dust storm image.

 Table 2

 Quantitative Evaluations of Different Methods using GMRVL (Hautiere et al. 2008), BRISQUE (Mittal et al. 2012) and PIQE (Venkatanath et al. 2015)

Metrics Metrics	Image Enhancement		Physical Prior-based		Supervised Learning		Semi-Supervised Learning		Unsupervised Learning	
	AL	CLAHE	DCP	NLP	MSCNN	EPDN	SSDN	DADN	ZID	Our
$ \begin{array}{c} \text{GMRVL} \uparrow \\ \text{BRISQUE} \downarrow \\ \text{PIQE} \downarrow \end{array} $	2.280 34.100 44.541	2.071 28.757 45.004	2.715 38.739 47.544	1.952 38.051 46.726	1.902 37.955 46.297	1.529 25.750 37.253	1.923 26.026 <u>33.351</u>	1.236 28.693 35.091	1.881 25.867 38.952	1.856 24.627 27.834

Note. ↑ indicates that the larger, the better. ↓ indicates that the smaller, the better. The bold and underlined values indicate the best and second-best results.

4.4. Ablation Studies

To evaluate the effectiveness of each configuration in our model, we compare the following models: (1) "BN": using the batch normalization in the entire networks; (2) "IN": using the instance normalization in the entire networks; (3) "LLS": using only one long skip connection that connects low-level stage, i.e., the encoder block behind the feature extraction and the generator block before the final output one; (4) "CNN": using the vanilla convolutional layers, i.e., removing the identity shortcut connections in each ResNet blocks; (5) CycleGAN: using CycleGAN framework, i.e., excluding the dust storm encoder E_{ds} . All of the above models are trained for ten epochs for fast evaluations.

As shown in Table 3, both of the results of models "BN" and "IN" reduce the performances of dust storm removal compared with our model with SSN. Furthermore, it verifies that different layers of the networks behave differently; thus, the normalization should also be changed at each layer. Moreover, they

would generate "spot" artifacts as labeled by the red rectangle illustrated in Figures 8(b) and (c). The model U-Net uses skip connections between the encoders and generators at each stage. Compared with the "LLS" where skip connections are only used at the low-level stage, U-Net shaped network increases GMRVL by 0.157, and it significantly reduces BRISQUE and PIQE by 1.179 and 8.226, respectively. These are because the skip-connections at different levels preserve the generators' structure (low-level features) and semantic (high-level features) information. The model "CNN" replaces the ResNet blocks with two stacked conventional convolutional layers. However, it reduces by 0.083 on GMRVL and increases by 0.769 on BRISQUE and 0.403 on PIQE. We further compare the DRL with the CycleGAN, another widely used framework for unsupervised learning. As illustrated in Table 3, the DRL framework contributes 0.453 improvements on GMRVL, 8.491 improvements on BRISQUE, and 10.509 improvements on PIQE, compared with the CycleGAN. These evaluations verify the effectiveness of using the DRL framework.

Figure 8. The visual results of the models in our ablation studies. As we can find that artifacts are generated in the results of the models of "BN", "IN", "LLS" and "CycleGAN", demonstrating the inefficiency of these models.

 Table 3

 Ablation Study: Evaluations on the SSN, U-Net Shape, ResNet Block and DRL (at 10th epoch)

Metrics	BN	IN	LLS	CNN	CycleGAN	Ours (SSN, U-Net, ResNet, DRL)
GMRVL ↑	1.434	1.662	1.529	1.603	1.233	1.686
BRISQUE ↓ PIQE ↓	26.466 30.081	25.863 29.640	26.792 36.229	26.382 28.406	34.104 38.512	25.613 28.003

5. Conclusion and Discussion

In this work, we propose an unsupervised Martian dust storm removal network via DRL. The network is formed with three parts, i.e., the forward translation, the backward reconstruction, and self-reconstruction, to achieve powerful representation learning. Additionally, we enforce the adversarial loss, cycleconsistency loss, and latent reconstruction loss to train the model.

Advantages. There are several significant advantages of our model. First, our dust storm removal model with DRL can be solely trained on the real-world unpaired dust storm and clear images. Thus, we can circumvent the expansive and timeconsuming data set collection and address the domain shift issue. Second, the model can implicitly learn the dust stormrelevant prior knowledge from the dust storm data sets, generating a high-quality image without relying on any handcrafted priors as training objects. Extensive experiments demonstrate that our model exhibits good generalization performances on real-world Martian dust storm removal.

Properties. It is worthy to note that our work is an image restoration model, not image enhancement. On the technical side, our model translates the Martian dust storm images into clear images. During this translation, the discriminator acts as a style classifier that can align the restored images and the clear images onto the same domain. In our data sets, the clear samples are collected from real, clear Martian images; as a result, our restored images are more looked like to be taken on a clear day on Mars. Because of this, unlike the goals of image enhancement algorithms that yield visually pleasant results with vivid color and high contrast, our goal is to restore the Martian dust storm image to a real, natural, and faithful one.

Beyond Martian dust storm removal. With the rapid development of space technology, numerous optical telescopes have been built (in outer space and on the ground), and massive amounts of astronomical image data have been captured. It enables us to develop data-driven-based approaches for astronomical image restoration. Unfortunately, it is impractical to simultaneously obtain the paired clear and degraded astronomical images. As discussed above, the advantages and properties of the proposed DRL network enable effective unsupervised learning, having huge potential for many other data-driven-based astronomy image restoration, where the paired unclear and clear images are unavailable.

Evaluation of the dust storm. We further propose a new approach for the Martian dust storm evaluations with respect to the level and spatial distribution in the scenes. Specifically, we compute the difference between the recovered image and the dust storm image by the following function:

$$\begin{aligned} \Delta'_{I} &= f_{g}[f_{c2g}[I^{c}]] - f_{g}[f_{c2g}[I^{d}]];\\ \Delta_{I} &= \Delta'_{I} + |\min(\Delta'_{I})|;\\ \overline{\Delta_{I}} &= \frac{\sum_{p \in \Delta_{I}} \Delta_{I}(p)}{H \times W}, \end{aligned} \tag{13}$$

where $f_{c2g}[\cdot]$ is the color-to-gray transform; $f_g[\cdot]$ is the Gaussian filter (window size = 7 and standard deviation = 3 in this paper). *H* and *W* represent the height and width of the image; *p* refers to the position of the pixel in the map. Δ_I is the map of the difference between the recovered image and the dust storm image; it also describes the distribution of the suspended dust particles. The Δ_I uses the middle equation in the Equation (13) to maintain that each pixel in the Δ_I map is > = 0. $\overline{\Delta_I}$ is the

Research in Astronomy and Astrophysics, 22:095020 (12pp), 2022 September

0.12 0.09 $\overline{\Delta_{I}} = 0.0898$ $\overline{\Delta}_{\rm I} = 0.0423$ 0.08 0.1 0.07 0.08 0.06 0.05 0.06 0.04 0.04 0.03 0.02 0.01 dust storm image Id recovered image Ic dust storm image Id recovered image Ic Δ_{I} Δ_{I} 0.14 $\overline{\Delta}_{\rm I} = 0.0505$ 0.08 $\overline{\Delta}_{\rm I} = 0.0367$ 0.12 0.07 0.1 0.06 0.05 0.08 0.04 0.06 0.03 0.04 0.02 0.02 0.01 recovered image I^c dust storm image I^d $\Delta_{\rm I}$ dust storm image I^c recovered image I^c Δ_{I}

Figure 9. Evaluation of the dust storm. Δ_I is the map of the difference between the recovered and dust storm images. As we can see, it can describe the spatial distribution of the dust. $\overline{\Delta_I}$ is the global average of Δ_I . Larger $\overline{\Delta_I}$ indicates more severe dust storms.

global average of Δ_i ; it reflects the level of the dust storm to some extent.

Figure 9 shows the visual results of Δ_I maps. As we can see, the Δ_I maps are consistent with the input images, indicating that the dust storm is heavier in the distant background (the sky or mountain) scenes while less in the foreground. The $\overline{\Delta_I}$ can be used as a new quantitative metric to measure the level of the dust storm in the scenes. As demonstrated in Figure 9, the left two samples got higher $\overline{\Delta_I}$ than the right ones, indicating that they encountered more severe dust storms.

Acknowledgments

This work was supported by the National Key R&D Program of China (No. 2021YFA1600504) and the National Natural Science Foundation of China (NSFC) (Nos. 11790305, 62132002 and 61922006).

References

- Alexandra Witze 2018, Massive Martian Dust storm Endangers NASA Rover, https://www.nature.com/articles/d41586-018-05425-1
- Ba, J. L., Kiros, J. R., & Hinton, G. E. 2016, arXiv:1607.06450
- Banfield, D., Spiga, A., Newman, C., et al. 2020, NatGe, 13, 190
- Berman, D., Treibitz, T., & Avidan, S. 2018, ITIP, 42, 720
- Chaffin, M., Kass, D., Aoki, S., et al. 2021, NatAs, 5, 1036
- Clarke, J. T. 2018, NatAs, 2, 114
- Collienne, P., Wolff, R., Gerndt, A., & Kuhlen, T. 2013, Workshop der GI-Fachgruppe VR/AR, Shaker Verlag, Aachen, 97
- Dominik Schmidt 2018, mars32k Dataset, https://dominikschmidt.xyz/ mars32k/
- Egan, W. G., & Foreman, K. M. 1971, Symposium-International Astronomical Union, Vol. 40 (Cambridge: Cambridge Univ. Press), 156
- Fattal, R. 2014, ACM Transactions on Graphics (TOG), 34
- Golts, A., Freedman, D., & Elad, M. 2019, ITIP, 29, 2692
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., et al. 2014, Proceedings of the Neural Information Processing Systems

- Hautiere, N., Tarel, J.-P., Aubert, D., & Dumont, E. 2008, Image Analysis & tereology, 27, 87
- He, K., Sun, J., & Tang, X. 2011, ITIP, 33, 2341
- He, K., Zhang, X., Ren, S., & Sun, J. 2016, in Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition,770
- Huang, X., & Belongie, S. 2017, in Proc. of the IEEE International Conf. on Computer Vision, 1501
- Ioffe, S., & Szegedy, C. 2015, ICML, PMLR, 448 Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. 2017, in Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, 1125
- Jobson, D. J., Rahman, Z.-u., & Woodell, G. A. 1997, ITIP, 6, 451
- Kingma, D. P., & Ba, J. 2015, in Proc. of the International Conf. on Learning Representations
- Leovy, C. 2001, Natur, 412, 245
- Li, B., Gou, Y., Liu, J. Z., et al. 2020, ITIP, 29, 8457
- Li, H., Cao, Y., Ding, M., & Zhuang, L. 2018, J. Beijing Univ. Aeronaut. onaut., 44, 444
- Li, L., Dong, Y., Ren, W., et al. 2019, ITIP, 29, 2766
- Lin, J., Xia, Y., Qin, T., Chen, Z., & Liu, T.-Y. 2018, in Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, 5524
- Liu, Y., Wu, X., Zhao, Y.-Y. S., et al. 2022, SciA, 8, eabn8555
- Mittal, A., Moorthy, A. K., & Bovik, A. C. 2012, ITIP, 21, 4695
- Narasimhan, S. G. 2000, in Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 1, 598
- Qu, Y., Chen, Y., Huang, J., & Xie, Y. 2019, in Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, 8160
- Ren, W., Pan, J., Zhang, H., Cao, X., & Yang, M.-H. 2020, Int. J. Comput. ision, 128, 240
- Reza, A. M. 2004, Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology, 38, 35
- Ronneberger, O., Fischer, P., & Brox, T. 2015, International Conference on Medical Image Computing and Computer-Assisted Intervention, 234
- Shao, W., Meng, T., Li, J., et al. 2019, in Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, 443
- Shao, Y., Li, L., Ren, W., Gao, C., & Sang, N. 2020, in Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, 2808
- Ulyanov, D., Vedaldi, A., & Lempitsky, V. 2017, in Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition
- Venkatanath, N., Praneeth, D., Bh, M. C., Channappayya, S. S., & Medasani, S. S. 2015, 2015 Twenty First National Conference on Communications (NCC), IEEE, 1
- Zhu, J.-Y., Zhang, R., Pathak, D., et al. 2017, Proceedings of the Neural Information Processing Systems
- Zhu, Q., Mai, J., & Shao, L. 2015, ITIP, 24, 3522