



SIDNet: A single image dedusting network with color cast correction

Jiayan Huang^{a,b}, Haiping Xu^c, Guanghai Liu^d, Chuansheng Wang^e, Zhongyi Hu^f,
Zuoyong Li^{a,*}

^a Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, College of Computer and Control Engineering, Minjiang University, Fuzhou 350121, China

^b College of Computer and Data Science, Fuzhou University, Fuzhou 350108, China

^c College of Mathematics and Data Science, Minjiang University, Fuzhou 350121, China

^d College of Computer Science and Engineering, Guangxi Normal University, Guilin 541004, China

^e Department of Systems Engineering, Automation and Industrial Informatics, Polytechnic University of Catalonia, Barcelona, Spain

^f Intelligent Information Systems Institute, Wenzhou University, Wenzhou 325035, China

ARTICLE INFO

Article history:

Received 7 October 2021

Revised 7 May 2022

Accepted 8 May 2022

Available online 11 May 2022

Keywords:

Image dedusting

Deep learning

Color cast correction

Dusty image synthesis

ABSTRACT

Dust degrades image content and causes image color cast, which negatively impacts on many high-level computer vision tasks. In this paper, we proposed a dedusting network with color cast correction for a single dusty image (SIDNet). The SIDNet contains several dust-aware representation extraction (DustAre) modules with the same structure. Each DustAre module contains two branches. The first branch encodes the input to estimate global veiling-light and local spatial information. The second branch generates a dust-aware map and fuses the global veiling-light, the local spatial information and the dust-aware map to generate the output. To further improve real dusty image dedusting performance, the SIDNet introduces a color cast correction scheme to our neural network. After considering that the average chromaticity values of a dusty image in CIELAB color space are usually larger than those of a clean (dust-free) image, the SIDNet defines a new loss function to better guide the network training. Additionally, we also construct a new synthetic dusty image dataset for network training, which additionally considers the scene depth relationship between real dusty image and dust-free image. Experiments on synthetic and real dusty images show that the SIDNet achieves better dedusting performance compared to state-of-the-art image restoration methods.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

In recent years, sand-dust storm weather has frequently been increasing, which negatively impacts on the ability of monitoring systems, such as license plate detection of an outdoor monitoring device and objective recognition technology of automatic driving. Images captured in sand-dust storm weather often depict color cast and lack contrast due to the scattering and absorption of light as it propagates through the dust. Dusty image input will make the other computer vision tasks difficult, such as image retrieve tasks [1], image classification [2], and image salient region detection [3]. Therefore, image dedusting as one pre-processing step of high-level computer vision tasks, has great significance to the field of computer vision. Fig. 1 shows several real dusty and clean (dust-free) images, as well as their corresponding RGB color histograms. From Fig. 1, it can be founded that the RGB colors of dusty im-

ages are usually concentrated in a certain subrange of their corresponding color histogram, and the distributions of R, G, B colors are relatively separated. On the contrary, the R, G, B colors of dust-free images are more evenly distributed in the whole range (i.e., [0, 255]), and the distributions of the three colors tend to overlap. Image dedusting aims to remove dust from a dusty image while maintaining image objects in their initial natural color. Existing image restoration methods can be divided into traditional color cast correction methods and deep learning-based methods.

Traditional color cast correction methods [4,5] usually first convert a dusty image into CIELAB color space with lightness and chromatic components separately. Then, they employ color cast correction on two chromatic components and detail enhancement on the lightness component to obtain the corresponding dust-free image. However, traditional image dedusting methods only focus on image color contrast adjustment and do not consider the relationship between dusty image scene depth and dust-free image scene depth. Therefore, the restoration results of traditional meth-

* Corresponding author.

E-mail address: fzulzytdq@126.com (Z. Li).

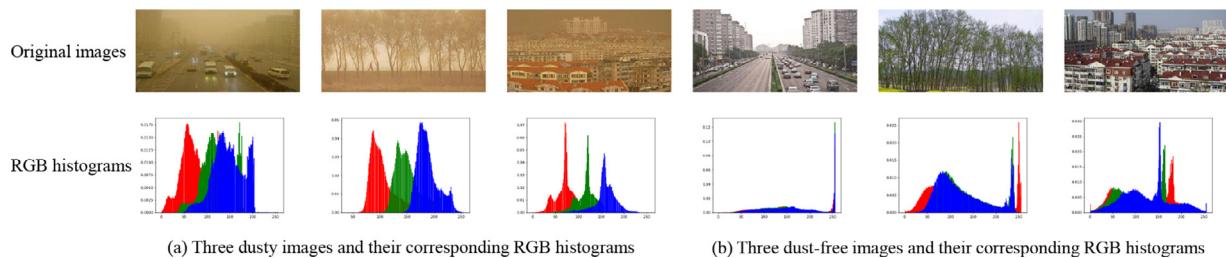


Fig. 1. Three pairs of real dusty images and dust-free images in similar scenes. The first row exhibits the six original images, and the second row exhibits their corresponding RGB color histograms.

ods often suffer from low color saturation and still with lots of residual dust.

Recently, deep learning technology has achieved effective application on the restoration of hazy images [6–8], rainy images [9,10], underwater images [11] and dusty images [12]. Deep learning-based methods usually design a convolution neural network for learning image features and training the network on a large-scale image dataset for generating the restoration image [13]. Since most existing networks are training on synthetic dusty image datasets, deep learning-based methods easily generate good effect on synthetic datasets. However, they show poor effect on real images [12]. Additionally, since there are few dusty image datasets for network training, most existing image dedusting research remains traditional.

Inspired by traditional color contrast correction methods [4,14,15] and existing deep learning-based image dehazing networks [16], we presented a network with color cast correction for single image dedusting (briefly called SIDNet). The SIDNet uses several same dust-aware representation extraction (DustAre) modules to downsample a dusty image to extract image features and then upsample them to obtain the final dedusting result. Each DustAre module contains two branches to estimate global veiling-light, local spatial information, and a dust-aware map, and then fuse them to generate the output. Furthermore, to improve the dedusting performance on real dusty images, the SIDNet introduces a color cast correction scheme. After comprehensively observing the differences between a dusty image and a dust-free image in their CIELAB color spaces, the SIDNet designs a new loss function to better guide the network training. In addition, due to the challenge of acquiring realistic dusty images, we take the relationship between real dusty image and dust-free image into account and construct a new dusty image synthetic dataset with both outdoor and indoor scenes for the network training. Quantitative and qualitative experimental results on a series of synthetic and real dusty images show that the SIDNet realized effect image dedusting compared to several state-of-the-art image restoration methods.

Overall, the main contributions of this paper are as follows:

- (1) Inspired by image dehazing, we presented a network with color cast correction for image dedusting (SIDNet).
- (2) The proposed SIDNet embeds several dust-aware representation extraction (DustAre) modules with the same structure, which is used to estimate and fuse the global veiling-light, the local spatial information, and the dust-aware map for improving image dedusting effect.
- (3) Considering that most existing deep learning-based image restoration methods have achieved good effect on synthetic images and poor effect on real images, as well as for avoiding network being excessively affected by real ground truth images with color cast, the SIDNet introduces a color cast correction scheme for those ground truth images from training dataset to meet both of synthetic and real dusty image dedusting performances.

- (4) After observing that dusty image usually has larger a^* and b^* average chromaticity values than the dust-free image in CIELAB color space, the proposed network designs a new loss function to better guide network training.
- (5) To enrich dusty image datasets for network training, we considered the relationship between dusty image scene depth and dust-free image scene depth, and constructed a synthetic dusty image dataset with both outdoor and indoor scenes for deep learning-based image dedusting networks training.
- (6) Experimental results on a series of synthetic dusty images and real dusty images show that the proposed SIDNet achieves a better image dedusting effect than several state-of-the-art image restoration methods, especially on real dusty images.

The rest of this paper is organized as follows. [Section 2](#) describes the related works. [Section 3](#) introduces the theory and implementation of the proposed method. [Section 4](#) analyzes the experiment results, and [Section 5](#) concludes this paper.

2. Related works

2.1. Image formation model learning

Similar to hazy images, the influence of dust on an image also depends on image scene depth. Therefore, the relationship of image scene depth between dusty and dust-free images should be considered when generating a dusty image. Instead of using the hazy image formation model [17,18] directly to depict that of a dusty image, we depict the formation of a dusty image according to its characteristics as,

$$I(x) = J(x)d(x)l(x) + C(1 - d(x)l(x)), \quad (1)$$

where I and J denote a dusty and a dust-free images, respectively. The symbol C denotes a dust color map, i.e., global veiling-light. The symbol l denotes the image scene depth map obtained by the pre-trained monocular image scene depth estimation method Monodepth2 [19]. d denotes the relationship between dusty image scene depth and dust-free image scene depth (briefly called dust-code). To estimate $d(x)$ accurately, we generate its value by the image scene depth estimation of 20 real dusty images and their corresponding dust-free images obtained by Photoshop in this paper.

Specifically, according to the atmospheric scattering model [17,18], the relationship between scene depth and transmittance is shown in Eq. (2),

$$t(x) = e^{-\beta l(x)}, \quad (2)$$

where $l(x)$ is the scene depth and $t(x)$ is the transmittance, which describes the proportion of atmospheric light penetrating the aerosol suspended in the air to reach the scene. The higher the transmittance, the less the scene is affected by dust. According to Eq. (2), the impact of dust on scene clarity is logarithmic and inversely proportional to $l(x)$ that is, the deeper the scene depth,

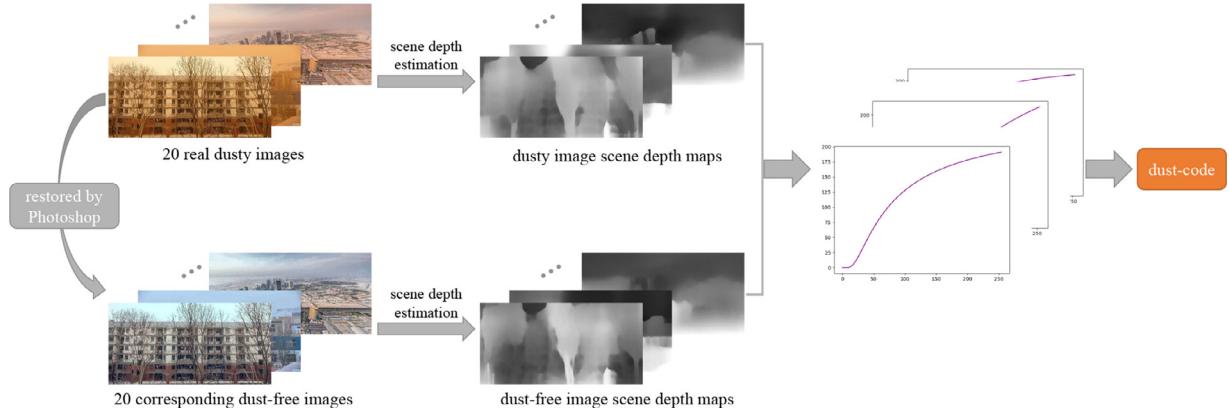


Fig. 2. The flow chart of dust-code estimation.

the smaller the transmittance. Furthermore, when the pixel value of the depth map is smaller, it means that the depth of field of the scene is deeper, which is just opposite to the representation of the transmittance map.

At the same time, since Monodepth2 [19] ignores the influence of bad weather on scene visibility during training, the output depth map cannot be used directly to synthesize realistic dusty images. The atmospheric scattering model indicates that there is an exponential relationship between the scene visibility and the scene depth, so the impact of scene visibility on the estimation scene depth result of Monodepth2 is also exponential relationship. Since dusty images have a significant impact on the scene visibility [20], we use Eq. (3) to estimate the relationship between the scene depth of clean image and that of dusty image based on mathematical statistics, which is equivalent to estimating the impact of dust on the scene visibility.

$$y = a \times e^{b/x}, \quad (3)$$

where a and b are the fitting parameters, and they are two arrays with the size of 1×255 . x and y represent the scene depth information of the clean image and the dusty image, respectively. Additionally, to estimate the relationship between the scene depth information of clean image and that of dusty image depth, we used the least square method, i.e., Eq. (3), which need two arrays a and b to find the relationship between two variables x and y . Moreover, although there are some similarities, the scene depth information relationships of different clean-dusty image pairs are not completely consistent in practice. Therefore, to reflect the general level of dust distribution for more relatively reasonable and natural dusty image synthesis, we average all a and b generated from 20 clean-dusty image pairs, and the resulting two arrays with the size of 1×255 are named dust-code d in Eq. (1). Fig. 2 gives the flow chart of dust-code estimation.

2.2. Traditional image color cast detection and correction

Color cast detection. The method [4] analyzes the characteristics of an image color distortion in CIELAB color space. CIELAB color space contains three components, that is, lightness component L ranging from 0 (dark) to 100 (light), a^* color component ranging from -127 (green) to 128 (red), and b^* color component range from -127 (blue) to 128 (yellow). The method further introduces the concept of equivalent circle, which takes average chromaticity values (d_{a^*} , d_{b^*}) as the circle center coordinate, and chromaticity center distance M as the radius. Furthermore, it uses the ratio of a^* , b^* jointed average chromaticity value D to chromaticity center distance M (i.e., color cast factor K) to measure the degree of the color cast of an image. The definitions of D , M , and K in

CIELAB color space are as follows,

$$D = \sqrt{d_{a^*}^2 + d_{b^*}^2}, \quad (4)$$

$$M = \sqrt{m_{a^*}^2 + m_{b^*}^2}, \quad (5)$$

$$K = \frac{D}{M}, \quad (6)$$

where $d_{a^*} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W a^*(i, j)$ and $m_{a^*} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (a^*(i, j) - d_{a^*})^2$ denote the average chromaticity value and chromaticity center distance of a^* component, respectively. $d_{b^*} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W b^*(i, j)$ and $m_{b^*} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (b^*(i, j) - d_{b^*})^2$ corresponding to that of b^* component. Generally speaking, an image is judged to have a color cast when the color cast factor $K \geq 1.5$.

Color cast correction. Based on the equivalent circle idea, the new corrected \hat{a}^* , \hat{b}^* components are generated by,

$$\hat{a}^* = a^* - d_{a^*}, \quad (7)$$

$$\hat{b}^* = b^* - d_{b^*}. \quad (8)$$

This technology can correct most images with different color casts. Fig. 3 gives a real dusty image and the restoration result according to Eqs. (7) and (8), and their corresponding equivalent circles. After color cast correction, the average chromaticity values of a^* , b^* components, and color cast factor K are obviously smaller, shown in Fig. 3(b) and (d). However, such color cast correction technologies exist a common problem that the color saturation of the restoration results are still relatively low, shown in Fig. 3(c).

2.3. Deep learning-based image restoration

Recently, convolution neural networks (CNNs) have been widely and successfully applied in restoring an image under bad weather. Owing to the available and publicly related synthetic datasets for network training [11,13], there are many deep learning-based methods focusing on hazy images [6–8], rainy images [9,10] and underwater images restoration [11]. Deep learning-based image restoration networks can be divided into two categories. One category network automatically estimates a transmission map and the atmospheric light value. Following this, it generates image dehazing results according to the atmospheric scattering model, such as [21]. The other category is end-to-end networks [7,16], which directly generate the image dehazing result by network learning and do not rely on the atmospheric scattering model. However, due to the lack of dusty image datasets, image dedusting technology

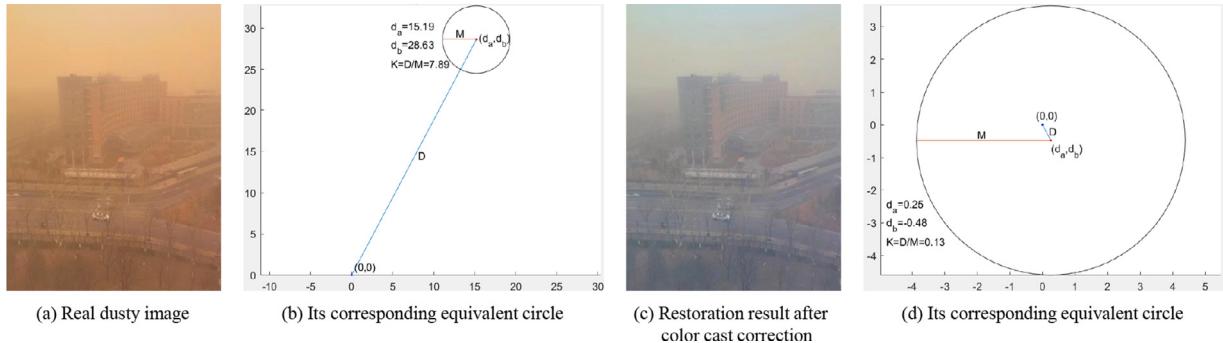


Fig. 3. Color cast correction on a real dusty image.

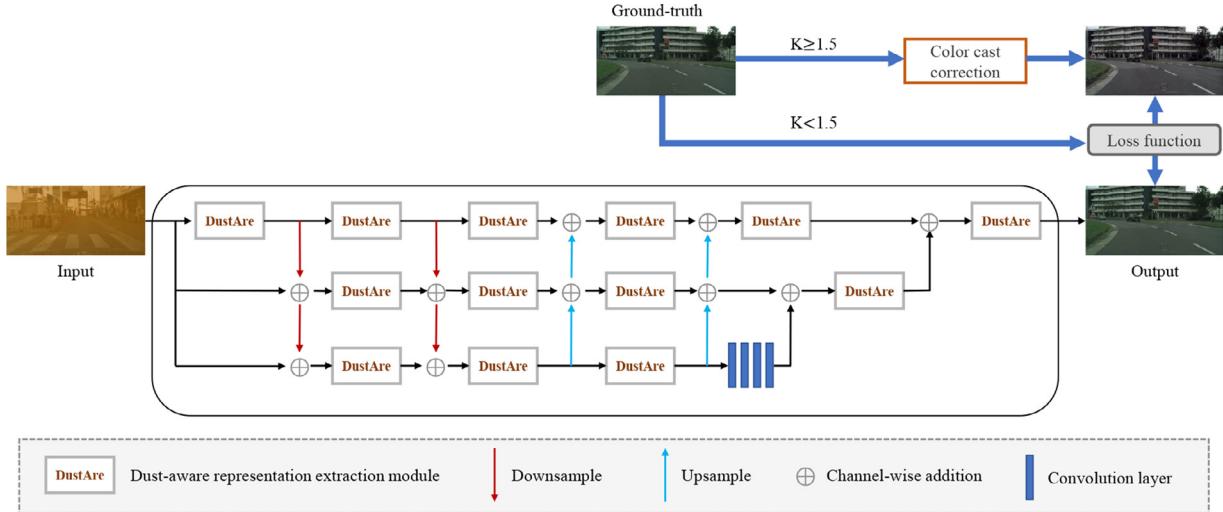


Fig. 4. Network structure of the proposed SIDNet.

still stays in various traditional color correction ways [4,5], whose restoration image saturation is still not enough. Furthermore, most of the existing deep learning-based methods have a common problem that they generate good effect on a synthetic dataset. However, they show poor effect on real scenes. Inspired by the wide application of CNNs, we presented a network with color cast correction for single image dedusting and constructed a new dusty image dataset for the network training.

3. The proposed method

To alleviate the common problem mentioned above, we presented an effective image dedusting network with color cast correction (SIDNet). The following subsection first introduces the network structure of the SIDNet, and then describes its important details, including dust-aware representation extraction (DustAre) module, color cast correction, and loss function.

3.1. Network structure

The proposed network uses several dust-aware representation extraction (DustAre) modules to downsample a dusty image for extracting image feature maps and then upsample the image feature maps to generate the final dedusting result. As shown in Fig. 4, the height and width of the network are three and six, respectively. Specifically, we use x_r^i , y_r^i to denote the input and output of the i -th DustAre module in the r -th row of the network, respectively. Then, during the downsample processing, the input of the $(i+1)$ -th DustAre module in the second and the third row can be

expressed as,

$$x_2^{i+1} = (\downarrow y_1^i) \oplus x_2^i, \quad (9)$$

$$x_3^{i+1} = (\downarrow\downarrow y_1^i) \oplus (\downarrow y_2^i) \oplus x_3^i. \quad (10)$$

Similarly, during the upsample processing, the input of the $(i+1)$ -th DustAre module in the second and the first row can be expressed as,

$$x_2^{i+1} = (\uparrow y_3^i) \oplus x_2^i, \quad (11)$$

$$x_1^{i+1} = (\uparrow\uparrow y_3^i) \oplus (\uparrow y_2^i) \oplus x_1^i, \quad (12)$$

where the operator \downarrow and \uparrow denote a downsample and a upsample operation, respectively.

3.2. Dust-aware representation extraction module

In order to remove the dust from an image while preserving image details, estimations of both global veiling-light and local spectral information are important. Inspired by the existing dehazing network [16], we use the dust-aware representation extraction module to extract the representations of them, as shown in Fig. 5. The dust-aware representation extraction module mainly contains two branches. The first branch uses a $1 \times 1 \times 2$ matrix to encode the input and generate the global veiling-light representation (i.e., β^V, γ^V), and a $H \times W \times 2$ matrix to generate the local spectral information representation (i.e., β^S, γ^S). The second branch is used to normalize two pair representations with the input x according

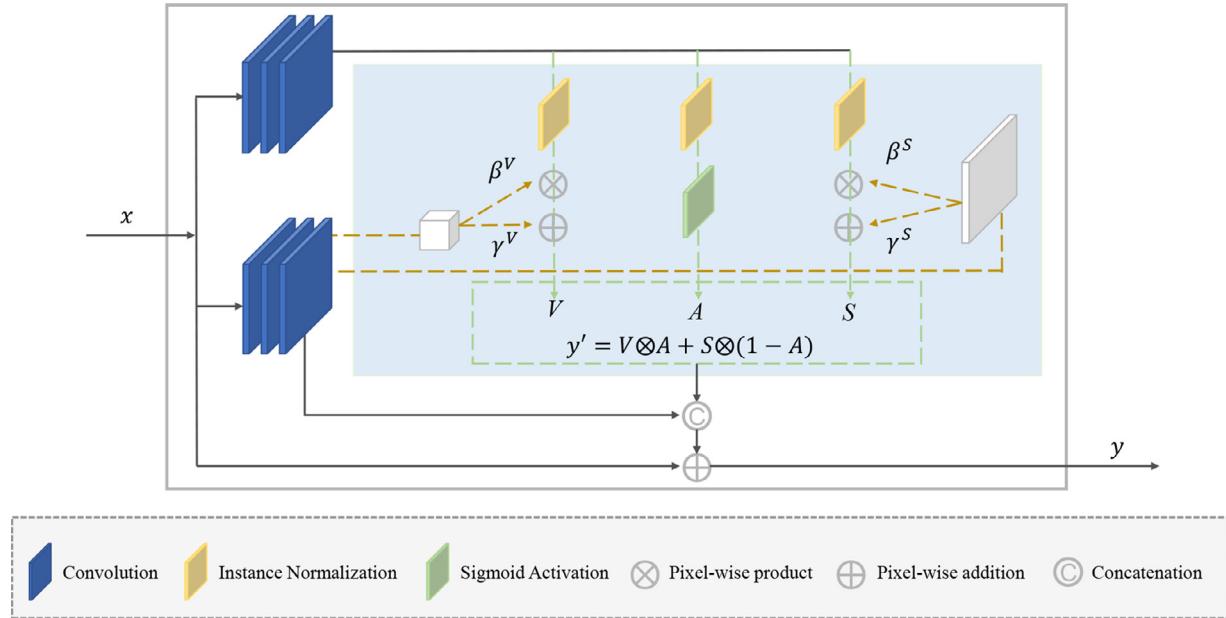


Fig. 5. Dust-aware representation extraction (DustAre) module.

to Eqs. (13) and (14) (defined as V and S , respectively), and generate a dust-aware map (A) by going through an instance normalization layer followed by a sigmoid layer at the same time. Following this, it fuses V and S based on A to generate the output y of the DustAre module. In detail, for the input x of the DustAre module, V , S and output y are expressed as,

$$V = \beta^V \frac{x - \mu}{\sigma} + \gamma^V, \quad (13)$$

$$S = \beta^S \frac{x - \mu}{\sigma} + \gamma^S, \quad (14)$$

$$y = V \otimes A + S \otimes (1 - A), \quad (15)$$

where μ and σ are denote the mean and standard deviation of x , respectively.

3.3. Color cast correction

In fact, most of the existing deep learning-based image restoration methods have a common problem that they achieve good effect on synthetic datasets while showing poor effect on real images. One of the main reasons is that the final image restoration results are overly dependent on the training dataset and even produce unnatural colors due to the color cast of the training dataset itself. On the one hand, the training set in the paper is synthesized based on two standard image datasets with scene depth maps (the outdoor clean image dataset Cityscape [22] and the indoor clean image dataset NYU-depth [23]). The Cityscape is captured from different urban scenes during driving with low brightness, and the NYU-depth is captured from indoor scenes that some images of this dataset have color cast under the influence of light. On the other hand, since images captured in sand-dust storm weather are often greatly affected by image scene brightness, improving the brightness of dusty images is also an essential step of the image dedusting task. Therefore, in order to make the model better applied to real dusty scenes with higher brightness and more natural scene color, the SIDNet uses color cast detection and correction for ground-truth images from the training dataset. Additionally, due to the problem of single scene of the two datasets, this pre-processing operation can enrich the scenes at a certain to adapt

to more realistic dusty scenes. Specifically, it thinks those ground-truth images with color cast factor $K \geq 1.5$ are more likely to have color cast themselves. Therefore, we correct a^* , b^* components of these images in CIELAB color space, i.e., final training ground truth image \hat{y} is generated by the following expression,

$$\hat{y} = \begin{cases} cor(\hat{y}), & K \geq 1.5 \\ \hat{y}, & \text{otherwise} \end{cases}, \quad (16)$$

where \hat{y} denotes an original ground-truth image, cor denotes the color cast correction operation mentioned in Section 2.2. Fig. 6 gives comparisons of original training ground truth images and their corresponding images after color correction. It can be observed that the brightness of the adjusted images is significantly improved, and the colors are more natural.

3.4. Loss function

The proposed SIDNet defines its loss function as,

$$\mathcal{L} = \lambda_{ab}\mathcal{L}_{ab} + \lambda_{L1}\mathcal{L}_{L1} + \lambda_{per}\mathcal{L}_{per}, \quad (17)$$

where λ_{ab} , λ_{L1} and λ_{per} are set to 0.3, 1.2 and 0.4, respectively. The first item \mathcal{L}_{ab} stems from the observation of dusty images and dust-free images. Due to a dusty image having serious color casts, the average a^* , b^* chromaticity values of a dusty image in CIELAB color space are usually larger than that of a dust-free image. Fig. 7 exhibits two similar real outdoor scenes with and without dust as samples to validate this observation. Fig. 7(a)-(d) show a real dusty image, its a^* component in CIELAB color space, its b^* component in CIELAB, and its equivalent circle, respectively. Similarly, Fig. 7(e)-(h) show a real dust-free image and its corresponding a^* component, b^* component, and equivalent circle, respectively. From Fig. 7, one can be observed that the a^* and b^* components of a dusty image are blurred and brighter than those of a dust-free image, which indicates the average a^* , b^* chromaticity values of a dusty image are larger than those of a dust-free image. This is consistent with the data given in Fig. 7(d) and (h). To further validate this observation, we calculate the a^* and b^* average chromaticity values on 50 real dusty and dust-free images. Experimental data shows that 50 real dusty images have an average a^* , b^* chromaticity values of 13.53 and 35.69, respectively. 50 dust-free images have an average

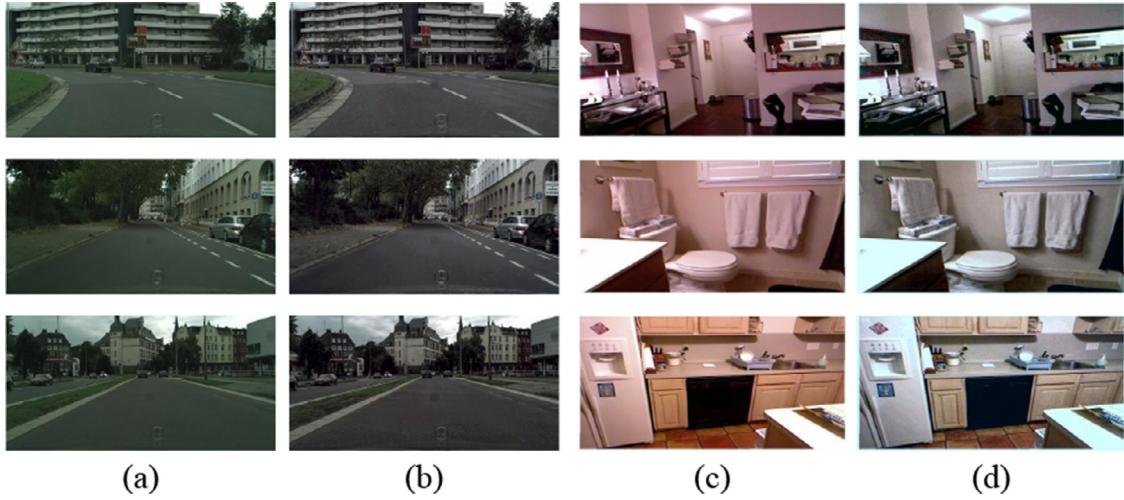


Fig. 6. Comparisons of original training ground-truth images and their corresponding images after color correction. (a) three original ground-truth images from Cityscapes, (b) after color cast correction corresponding to (a); (c) three original ground-truth images from NYU-depth, (d) after color cast corresponding to (c).

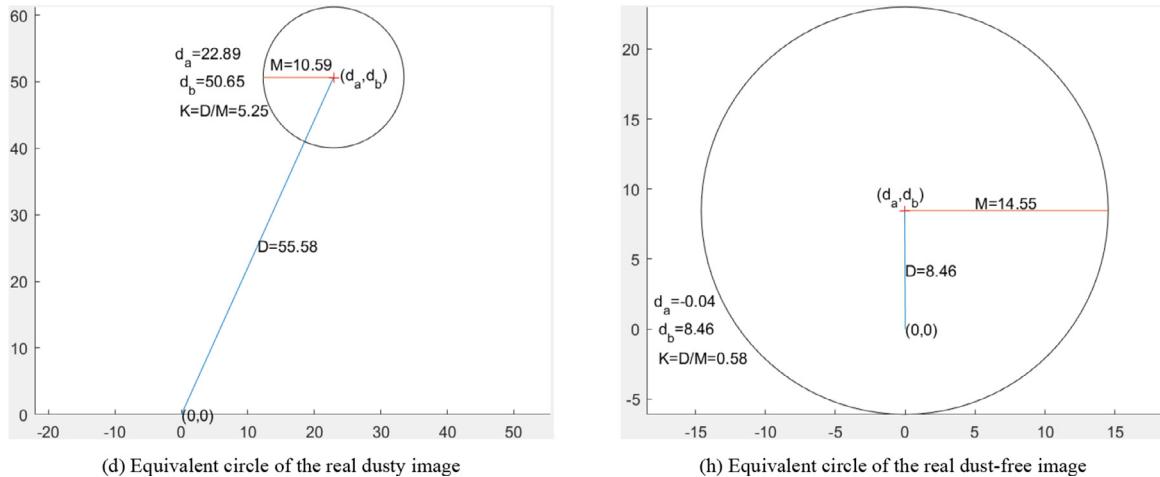


Fig. 7. The a^* , b^* components on a pair of similar scenes with and without dust.

a^* , b^* chromaticity values of 0.28 and 2.99, respectively. Following these data, one can conclude that a dusty image usually has large average a^* , b^* chromaticity values in CIELAB color space. Therefore, we deduce that the smaller the average a^* , b^* chromaticity values of an image dedusting result, the better the image dedusting effect. Motivated by the observations above mentioned, we designed the first loss item \mathcal{L}_{ab} as,

$$\mathcal{L}_{ab} = \left(\frac{1}{N} \sum_{i=1}^N y_a(i) - \frac{1}{N} \sum_{i=1}^N \hat{y}_a(i) \right)^2 + \left(\frac{1}{N} \sum_{i=1}^N y_b(i) - \frac{1}{N} \sum_{i=1}^N \hat{y}_b(i) \right)^2, \quad (18)$$

where N is the total number of image pixels. y_a and y_b denote the a^* and b^* color components in CIELAB color space of dedusting result y , respectively. \hat{y}_a and \hat{y}_b denote the a^* and b^* color components in CIELAB color space of corresponding dust-free image (i.e., ground-truth \hat{y}) after color cast correction \hat{y} , respectively.

The second item is smooth $L1$ loss termed as \mathcal{L}_{L1} . The \mathcal{L}_{L1} can prevent potential gradient explosion [24], which is defined as,

$$\mathcal{L}_{L1} = \sum_{c=1}^3 \sum_{i=1}^N \alpha(\hat{y}_c(i) - y_c(i)), \quad (19)$$

$$\alpha(k) = \begin{cases} 0.5k^2, & \text{if } |k| < 1 \\ |k| - 0.5, & \text{otherwise} \end{cases}. \quad (20)$$

The final item, perceptual loss \mathcal{L}_{per} , is used to encourage dedusting image to be represented with a similar feature in the backbone network, such as VGG19 pre-trained [25] in this work, and its definition is,

$$\mathcal{L}_{per} = \sum_{i=1}^3 \frac{1}{C_i H_i W_i} \| \phi(y) - \phi(\hat{y}) \|, \quad (21)$$

where H , W , and C denote height, weight, and channel of the feature map in the i -th layer of the network, respectively. $\phi_i(y)$ and $\phi_i(\hat{y})$ denote the output feature map in the i -th layer of y and \hat{y} ,

Table 1

The training details of the compared methods.

| Methods | LPNet [9] | RGNet [10] | AODNet [7] | FFANet [8] | Wang et al. [4] | HardGAN [16] | FFNet [12] | SIDNet |
|---------------|------------|------------|------------|------------|-----------------|--------------|------------|-----------|
| Frame | TensorFlow | TensorFlow | PyTorch | PyTorch | MatLab | PyTorch | PyTorch | PyTorch |
| Task | deraining | deraining | dehazing | dehazing | dedusting | dehazing | dedusting | dedusting |
| Epochs | 200 | 200 | 200 | 200 | / | 200 | 200 | 200 |
| Learning_rate | 1e-3 | 1e-3 | 1e-4 | 1e-4 | / | 1e-4 | 1e-4 | 1e-4 |
| Batch_size | 8 | 8 | 8 | 8 | / | 8 | 8 | 8 |
| Image_size | 512 × 256 | 512 × 256 | 512 × 256 | 512 × 256 | 512 × 256 | 512 × 256 | 512 × 256 | 512 × 256 |

respectively. y and \hat{y} denote dedusting result and corresponding dust-free image, respectively.

4. Experiments

To quantitative and qualitative evaluation image dedusting performance, we compared the SIDNet with seven state-of-the-art image restoration methods, i.e., LPNet [9], RGNet [10], AOD-Net [7], FFANet [8], Wang et al. [4], HardGAN [16], FFNet [12]. Among the seven methods, where LPNet, RGNet focus on image deraining, AOD-Net, FFANet, and HardGAN focus on image dehazing, while the other two methods concentrate on image dedusting. It is worth noting that all the compared methods are deep-learning based, except the method of Wang et al.. Specifically, the training details of all the compared methods are shown in Table 1, where the learning rate of each method is following their original setting.

This section firstly describes the dusty image datasets for network training and evaluation metrics for image dedusting capability. Then, it presents quantitative and qualitative evaluation results of the eight methods on synthetic and real dusty images. Next, it verifies the significance of image dedusting on image edge extraction and local keypoints matching. Finally, it analyzes the influence of several important designs in the proposed SIDNet, including DustAre module, color cast correction, and loss function item.

4.1. Datasets and evaluation metrics

There are few publicly available dusty image datasets which is difficult to collect in practice. Synthetic datasets are widely used to train models in the image enhancement fields. For example, Zhang et al. [26] used Photoshop to assist in image deraining task, and Li et al. [27] selected 3D movie pre-processing to provide outdoor stereo image pairs for haze synthesis. Therefore, we also used our synthetic dusty image dataset for several network training. In detail, the outdoor dusty images are synthesized from the Cityscapes dataset [22], while the indoor dusty images are synthesized from the NYU-depth dataset [23]. The dusty image dataset is composed of three subsets: outdoor synthetic training set, indoor synthetic training set, and synthetic testing set. The outdoor and indoor synthetic training set contains 800 (200 original dust-free images and their 600 synthetic images dusty images) and 2352 (588 original dust-free images and their 1764 synthetic dusty images) images, respectively. The synthetic testing set is composed of 268 outdoor synthetic dusty images (67 original dust-free images and 201 synthetic dusty images) and 308 indoor synthetic dusty images (77 original dust-free images and 231 synthetic dusty images). In addition, to verify the effectiveness of the proposed SIDNet on the real dusty scene, we chose 55 representative real dusty images under different scenes as testing samples. Fig. 8 exhibits some example images from our synthetic dusty dataset and real dusty images. It can be observed that the RGB color histograms of synthetic dusty images conform to the distribution characteristics of real dusty images, i.e., the three-color distributions of a dusty image are relatively separated and concentrated in the centralized sub-interval of the histogram.

To quantitatively compare dedusting performances of different methods, we use PSNR (peak signal to noise ratio) and SSIM (structural similarity) [28] to evaluate dedusting results of synthetic dusty images with corresponding reference images and use spatial-spectral entropy-based quality (SSEQ) [29] and image entropy (E) [30] to evaluate dedusting results of real dusty images without corresponding reference images. Among all the metrics, higher PSNR, SSIM, and E values indicate better dedusting results and SSEQ value range between 0 (best) to 100 (worst). To make the correlation consistent with PSNR and SSIM metrics, we reverse the SSEQ value in this paper.

4.2. Quantitative evaluation

To make a quantitative comparison of different methods, Tables 2–4 list average PSNR, SSIM, and SSEQ values of different methods on synthetic and real dusty images. The best and second-best quantitative data of the three tables were marked in red and blue, respectively. Specifically, Tables 2 and 3 list average PSNR and SSIM values of outdoor and indoor synthetic dusty image dedusting results obtained by different methods, respectively. Table 4 lists SSEQ values of eight real dusty images dedusting results, as well as their corresponding average SSEQ values obtained by different methods.

Table 2 shows that the proposed SIDNet obtained the highest average PSNR (32.54) and the highest SSIM (0.9294) values on the dedusting results of outdoor synthetic dusty images. HardGAN and FFNet exhibit their effective dedusting performance with higher average PSNR and SSIM values than other reference methods. FFANet, AODNet, LPNet, and RGNet display intermediate dedusting performance with middle average PSNR and SSIM values. Wang et al. exhibit their poor dedusting performance with the lowest average PSNR and SSIM values. Table 3 demonstrates that the SIDNet achieves the highest average PSNR (32.82) and SSIM (0.9432) values on the dedusting results of indoor synthetic dusty images indicating better dedusting performances than other reference methods. HardGAN gains the second-highest PSNR and the second-highest SSIM values. FFNet achieves the third-highest PSNR and SSIM values, while other reference methods obtain lower PSNR and SSIM values showing the worse dedusting performances.

To further validate the performance of image dedusting on real dusty images, eight representative real dusty images were selected as testing samples. These images depict several common dusty scenes, such as suburbs, roads, city, and outside of a window. Quantitative performances in Table 4 exhibit that the proposed SIDNet obtains the highest SSEQ values on most of the images, excluding the fifth and the eighth images, and achieves the second-highest SSEQ value on the first image. Furthermore, the SIDNet exhibits the best dedusting performance on the real dusty images with the highest average SSEQ value of 69.66. Among the seven counterparts, AODNet, RGNet, and FFANet obtain the highest SSEQ values on the first, fifth and last image, respectively. The method of Wang et al. exhibits better dedusting performances with the second-highest SSEQ values on the third to seventh images and the

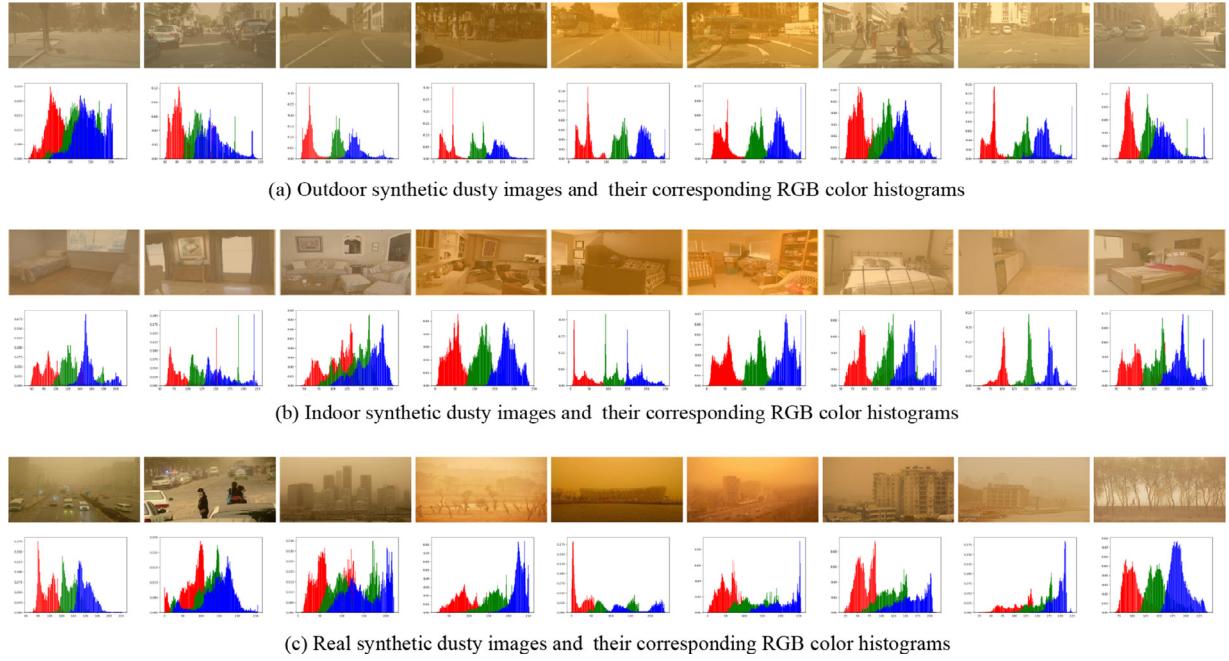


Fig. 8. Example images from our synthetic dusty image dataset and 9 real dusty images.

Table 2

Average PSNR and SSIM values of outdoor synthetic dusty image dedusting results obtained by different methods. The data marked in red and blue indicate the best and the second-best quantitative evaluation results, respectively.

| Methods | LPNet [9] | RGNet [10] | AODNet [7] | FFANet [8] | Wang et al. [4] | HardGAN [16] | FFNet [12] | SIDNet |
|---------|-----------|------------|------------|------------|-----------------|--------------|------------|--------|
| PSNR | 18.72 | 16.04 | 18.81 | 20.08 | 14.67 | 30.49 | 31.01 | 32.54 |
| SSIM | 0.8552 | 0.7825 | 0.7971 | 0.8440 | 0.7606 | 0.9272 | 0.9251 | 0.9294 |

Table 3

Average PSNR and SSIM values of indoor synthetic dusty image dedusting results obtained by different methods. The data marked in red and blue indicate the best and the second-best quantitative evaluation results, respectively.

| Methods | LPNet [9] | RGNet [10] | AODNet [7] | FFANet [8] | Wang et al. [4] | HardGAN [16] | FFNet [12] | SIDNet |
|---------|-----------|------------|------------|------------|-----------------|--------------|------------|--------|
| PSNR | 16.37 | 19.69 | 17.16 | 17.17 | 15.42 | 30.83 | 27.03 | 32.82 |
| SSIM | 0.8359 | 0.8439 | 0.8091 | 0.7894 | 0.7859 | 0.9381 | 0.9266 | 0.9432 |

Table 4

Average SSEQ values of real dusty image dedusting results obtained by different methods. The data marked in red and blue indicate the best and the second-best quantitative evaluation results, respectively.

| Image No. | LPNet [9] | RGNet [10] | AODNet [7] | FFANet [8] | Wang et al. [4] | HardGAN [16] | FFNet [12] | SIDNet |
|-----------|-----------|------------|------------|------------|-----------------|--------------|------------|--------|
| 1 | 55.80 | 53.11 | 63.06 | 54.21 | 53.15 | 51.90 | 53.10 | 58.64 |
| 2 | 67.73 | 73.43 | 80.63 | 67.93 | 75.05 | 74.02 | 85.38 | 87.78 |
| 3 | 58.56 | 65.58 | 53.72 | 67.80 | 71.70 | 56.56 | 53.72 | 74.48 |
| 4 | 63.63 | 64.84 | 68.07 | 62.24 | 68.13 | 66.61 | 66.70 | 68.49 |
| 5 | 58.27 | 66.84 | 51.03 | 61.74 | 65.39 | 60.81 | 63.80 | 60.76 |
| 6 | 61.93 | 67.96 | 65.12 | 65.24 | 74.05 | 66.46 | 66.01 | 75.69 |
| 7 | 59.92 | 69.27 | 67.74 | 68.54 | 69.54 | 67.91 | 67.30 | 70.69 |
| 8 | 60.48 | 61.74 | 45.61 | 69.53 | 61.84 | 55.34 | 61.87 | 60.78 |
| Average | 60.79 | 65.34 | 61.87 | 64.65 | 67.35 | 62.05 | 64.74 | 69.66 |

second-highest average SSEQ value. FFNet gains the second-highest SSEQ values on the second and the last images. LPNet exhibits the worst dedusting performance with the lowest average SSEQ value of 60.97.

4.3. Qualitative evaluation

To make a qualitative comparison of different methods, Figs. 9–11 exhibit visual dedusting results of the eight methods on several representative synthetic and real dusty images. Specifically, Figs. 9 and 10 exhibit the dedusting performance of different meth-

ods on six outdoor and indoor synthetic dusty images, respectively. Fig. 11 exhibits the dedusting performance of different methods on eight real dusty images.

From Fig. 9, it can be observed that the outdoor image dedusting results obtained by the proposed SIDNet are close to the ground-truth (i.e., dust-free images). HardGAN and FFNet obtain better dedusting performances than other reference methods. AODNet fails to remove dust for certain scenes. LPNet, RGNet and FFANet cause different degrees of color distortions. Wang et al. correct dusty image color. However, their dedusting results show low color saturation. For the indoor synthetic image dedusting results,

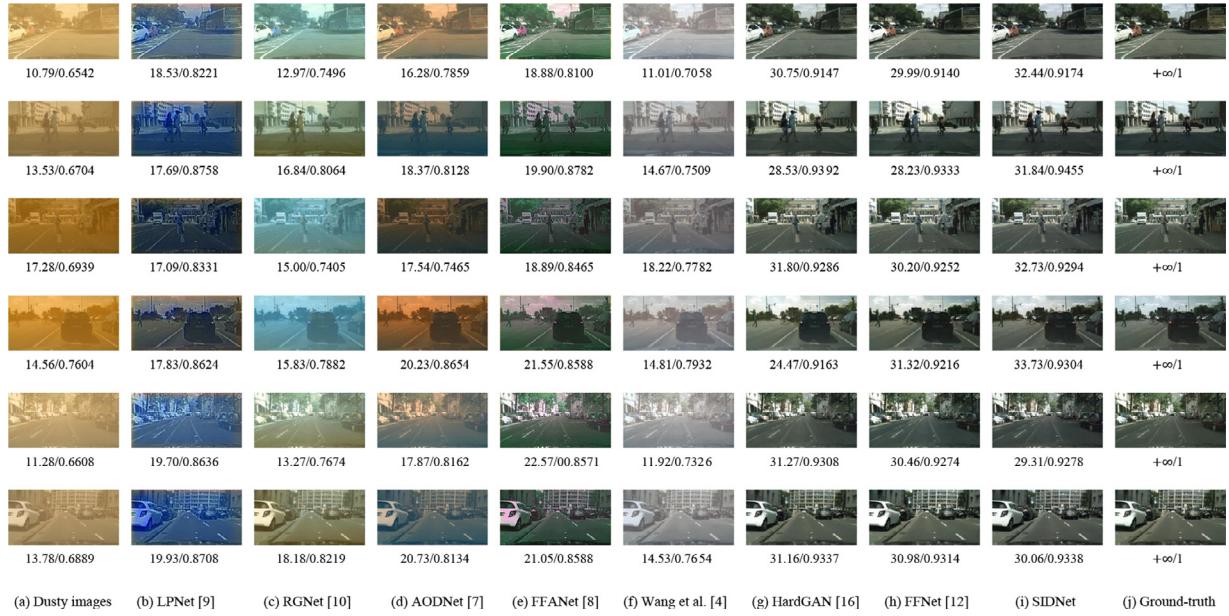


Fig. 9. Visual image dedusting results of different methods on outdoor synthetic dusty images.



Fig. 10. Visual image dedusting results of different methods on indoor synthetic dusty images.

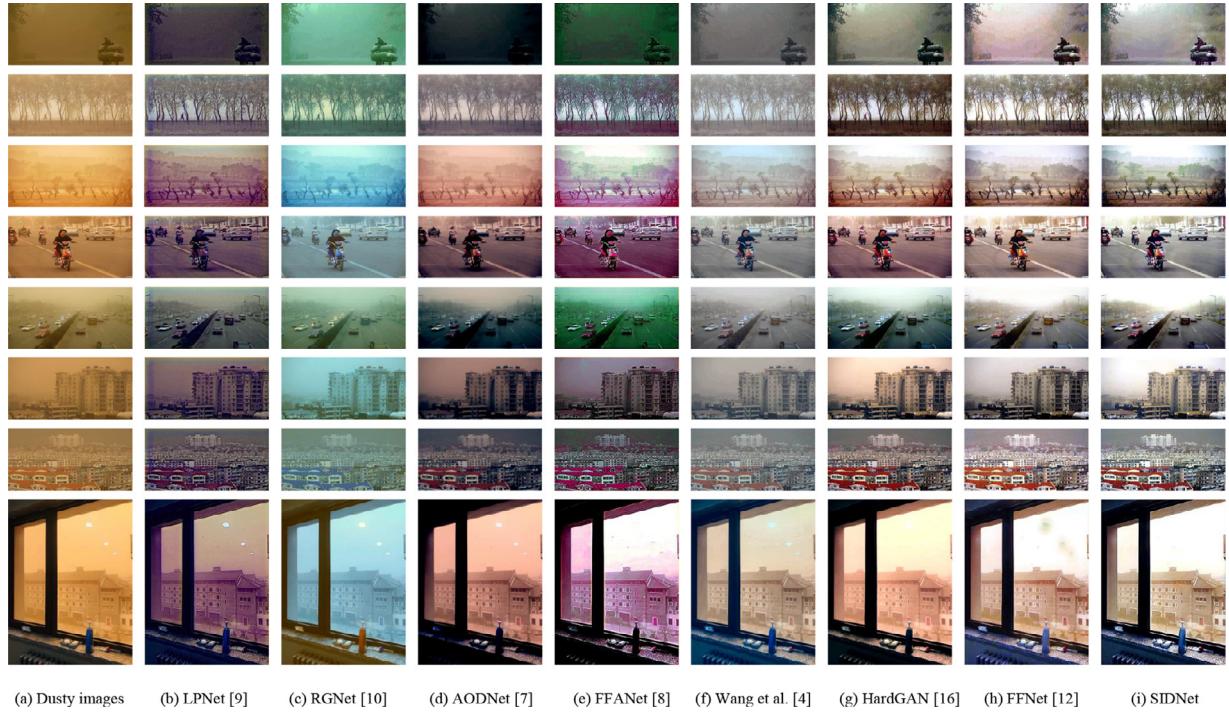
as shown in Fig. 10, the dedusting results obtained by the SIDNet are more similar to the ground-truth, indicating the best image dedusting performances. HardGAN and FFNet remove most of the dust effectively, and their dedusting results are also closed to the ground-truth. The dedusting results obtained by Wang et al. have low color saturation. LPNet, RGNet, and FFANet cause different color distortions. AODNet removes slight dust and shows low color saturation in some dedusting results, such as the last image.

To further validate image dedusting performance, we trained eight methods on synthetic dusty images and tested them on several real dusty images. Fig. 11 exhibits dedusting performances of different methods on eight representative real dusty images. These images contain different common dust scenes: forest with heavy dust-storm, multi-detailed trees, desert scene, road with short-distance persons, road with several cars, buildings under the sky, buildings with dense red roofs, and dust-storm outside the windows.

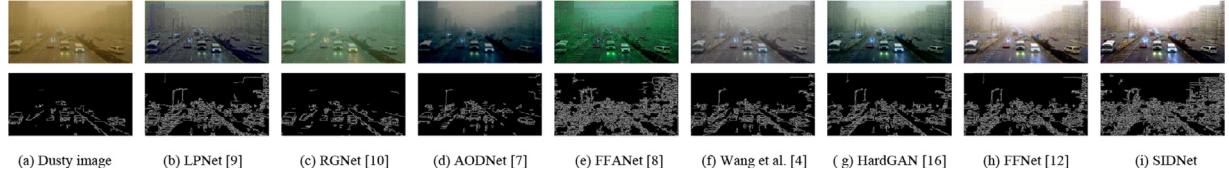
From Fig. 11, it can find that LPNet and AODNet remove slight dust, and dedusting results are darkish. RGNet and FFANet fail to remove most dust and cause serious color distortion. The method of Wang et al. removes most of the dust, however, the color of their dedusting results in low saturation. HardGAN and FFNet achieve better dedusting results than the former five methods. However, there are still dust residues on their dedusting results, such as the third, the fourth, and the last images. The SIDNet generates a good tradeoff between better image dedusting results and high color saturation.

4.4. Potential applications

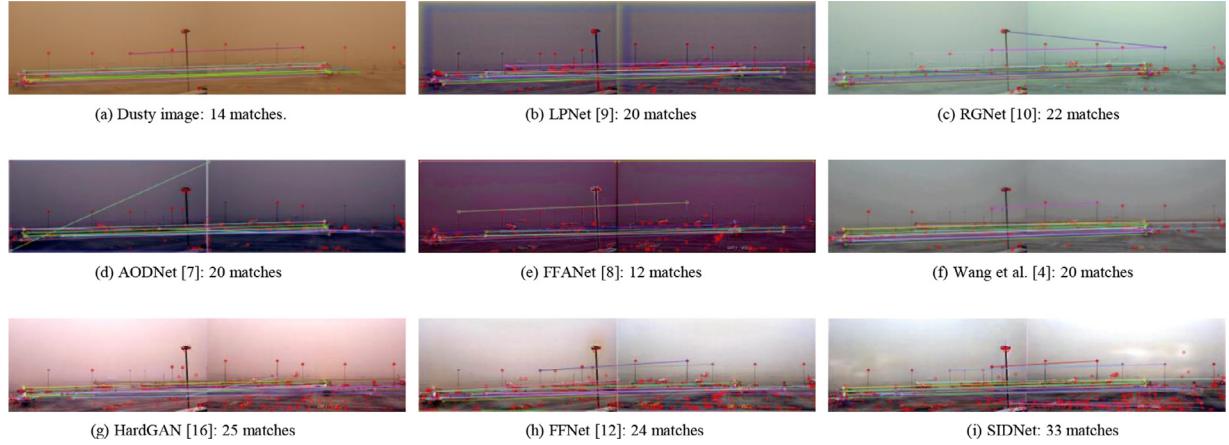
To validate the image dedusting could benefit other vision tasks, we performed two potential applications, namely image edge extraction and local keypoints matching.



(a) Dusty images (b) LPNet [9] (c) RGNet [10] (d) AODNet [7] (e) FFANet [8] (f) Wang et al. [4] (g) HardGAN [16] (h) FFNet [12] (i) SIDNet

Fig. 11. Visual image dedusting results of different methods on real dusty images.

(a) Dusty image (b) LPNet [9] (c) RGNet [10] (d) AODNet [7] (e) FFANet [8] (f) Wang et al. [4] (g) HardGAN [16] (h) FFNet [12] (i) SIDNet

Fig. 12. Edge extraction results on a real dedusted sample.**Fig. 13.** Number of matching points on a pair of similar scenes.

4.4.1. Image edge extraction

The color cast and low color contrast of dusty images make the edge extraction of dusty images difficult. To validate the significance of the SIDNet on image edge extraction performance, we analyzed the performances of image edge extraction on dedusting images of different methods. Compared with other counterparts, the edge details are increased in the dedusting results of the SIDNet as shown in Fig. 12. This validates that the SIDNet can significantly improve image edge extraction performance with more details.

4.4.2. Local keypoints matching

Local keypoints matching aims to find correspondences between two similar scenes. To validate the significance of the SIDNet on local keypoints matching, we use the SIFT (scale-invariant feature transform) operator [31] for a pair of similar real heavy dusty images (captured from the same airport), and their corresponding dedusting image pairs obtained by different methods. Matching results are exhibited in Fig. 13. One can see that the number of matching keypoints is obviously increased in the dedusting image



Fig. 14. The brightness comparisons of two real dusty image dedusted samples obtained by MSCNN module and DustAre module, respectively.

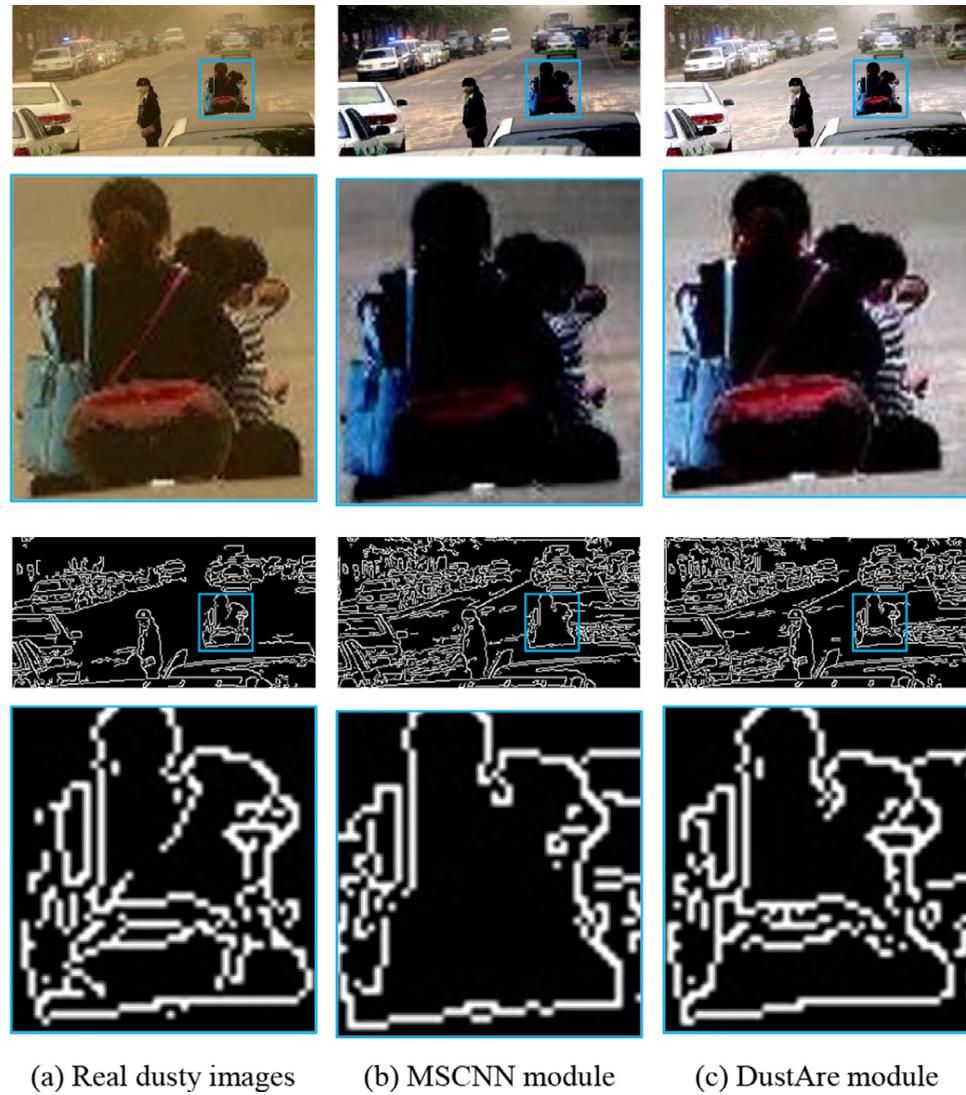


Fig. 15. Edge extraction effects of a real dusty image dedusted sample obtained by MSCNN module and DustAre module, respectively.

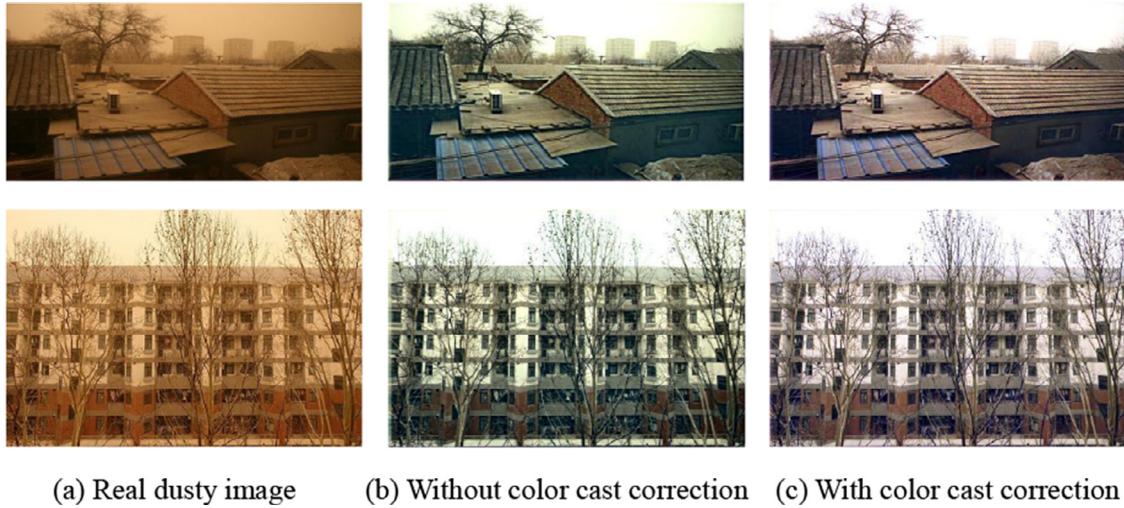
pairs. This confirms that the SIDNet can also recover local features of dusty images.

4.5. Discussions

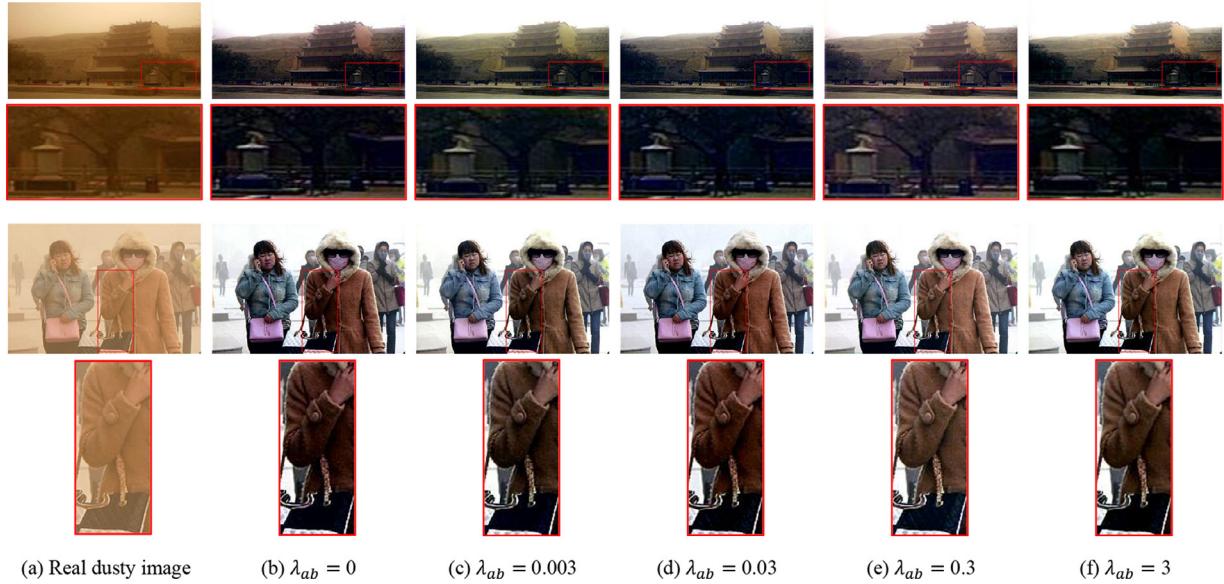
4.5.1. Analysis of DustAre module

To improve the performances of dedusting results in real dusty image, we introduce the Dust-aware representation extraction

(DustAre) module to the SIDNet, which is mainly used to estimation global veiling-light (V) and local spectral information (S). For the estimation of V , since our training dataset is mainly concentrated on urban scenes (Cityscapes dataset) and indoor scenes (NYU-depth dataset), the illumination information is quite different from the real dusty scenes. For example, both of Cityscapes and NYU-depth datasets basically do not include the sky area, but more sky areas represent higher brightness. Therefore, in order



(a) Real dusty image (b) Without color cast correction (c) With color cast correction

Fig. 16. Impact of color cast correction on the SIDNet dedusting effect.

(a) Real dusty image

(b) $\lambda_{ab} = 0$ (c) $\lambda_{ab} = 0.003$ (d) $\lambda_{ab} = 0.03$ (e) $\lambda_{ab} = 0.3$ (f) $\lambda_{ab} = 3$ **Fig. 17.** Qualitative comparison results of the SIDNet with different λ_{ab} 's weights.

to make our model more suitable for complex lighting conditions under natural sand-dust storm weather conditions, we introduce the DustAre module to estimate the global veiling-light (V). To prove the superiority of the DustAre module in the estimation of V , we compared it with other modules (such as Multi-Scale Convolutional Neural Network, MSCNN). The experimental results given in Fig. 14 demonstrate that the MSCNN module is difficult to adapt to scenes with large-area thick dust. On the contrary, the DustAre module has higher brightness on both thin and thick dust scenes.

For the estimation of S , since a dusty image is different from a general degraded image, it dramatically impacts the scene. Moreover, different dust colors have different effects on the semantic information of different objects. Therefore, in order to retain the local spatial information of the dedusted results, we introduce the DustAre module to estimate S . To prove the superiority of the DustAre module in the restoration of S , we compared it with MSCNN. The experimental results given in Fig. 15 demonstrate that the edge extraction result of the DustAre module can restore more image details than the MSCNN module.

Table 5

Impact of color cast correction on training label images of the SIDNet for image dedusting. The data marked in red indicate the best quantitative evaluation results.

| The SIDNet | SSEQ | E |
|-------------------------------|-------|------|
| Without color cast correction | 68.72 | 7.37 |
| With color cast correction | 69.70 | 7.45 |

4.5.2. Analysis of color cast correction strategy

To better remove dust from real dusty images and maintain more natural scene colors, the proposed SIDNet corrected the color of the ground-truth images of the training dataset by giving a certain threshold $K=1.5$. To validate the impact of color cast correction for ground-truth on real images dedusting, we tested on 20 real dusty images and exhibited two representative dedusting results obtained by the SIDNet without and with color cast correction, respectively. The quantitative and qualitative comparison results are given in Table 5 and Fig. 16. Experimental results show

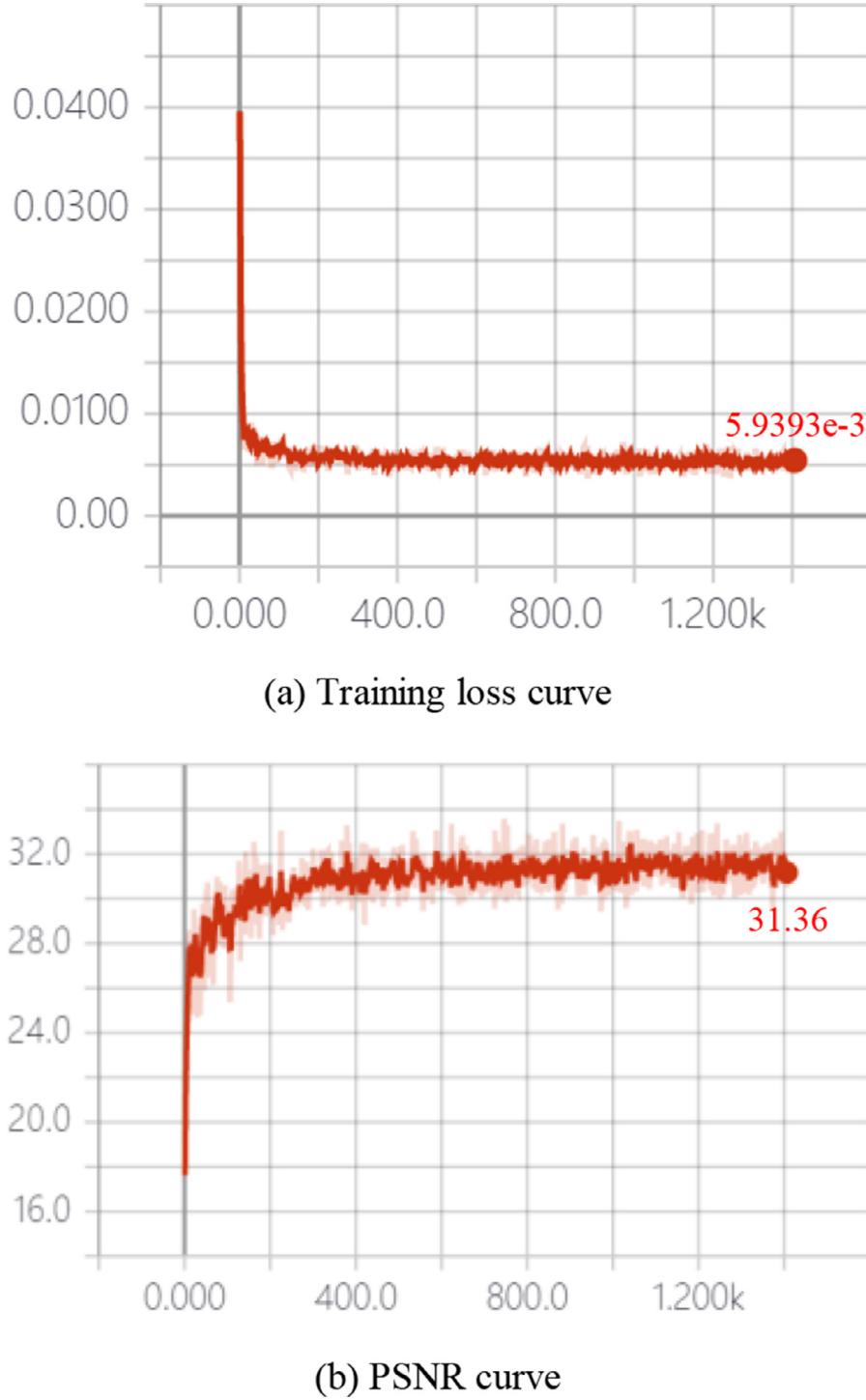


Fig. 18. Training error curve.

that the SIDNet with color cast correction for ground-truth images generates higher average SSEQ and E values and improves the image dedusting effect with more natural scene colors.

4.5.3. Analysis of \mathcal{L}_{ab} 's weight

Motivated by the observation that a dusty image usually has larger a^* and b^* average chromaticity values in CIELAB color space than that of a dust-free image, the SIDNet designs a new loss item \mathcal{L}_{ab} in Eq. (18). To validate the impact of \mathcal{L}_{ab} on the dedusting results, we followed the quantitative setting of 10 times difference in [32], and trained SIDNet with λ_{ab} as 0, 3×10^{-3} , 3×10^{-2} , 3×10^{-1}

and 3, respectively. Then, all models were evaluated on 20 real dusty images. The quantitative and qualitative comparison results are given in Table 6 and Fig. 17. As shown, when $\lambda_{ab} = 0.3$, the SIDNet generates higher average SSEQ and E values, improves the local dark shadow, and maintains the original high bright region without overexposure. However, others are getting too dark local shadow.

4.5.4. Analysis of training error curve

After analysis of \mathcal{L}_{ab} 's weight, we apply a more acceptable λ_{ab} value of 0.3 for our model training. The training loss curve is

Table 6

Quantitative comparison results of the SIDNet with different \mathcal{L}_{ab} 's weights. The data marked in red indicate the best quantitative evaluation results.

| the SIDNet | SSEQ | E |
|------------------------|-------|------|
| $\lambda_{ab} = 0$ | 67.16 | 6.97 |
| $\lambda_{ab} = 0.003$ | 68.34 | 7.03 |
| $\lambda_{ab} = 0.03$ | 68.33 | 7.06 |
| $\lambda_{ab} = 0.3$ | 69.70 | 7.45 |
| $\lambda_{ab} = 3$ | 65.34 | 6.69 |

shown in Fig. 18(a), and it can be observed that the SIDNet converges quickly and easily. Additionally, to demonstrate the efficiency of our model, we further give the PSNR curve during the training in Fig. 18(b). All experiments are performed on three NVIDIA Tesla P100 GPUs.

5. Conclusion

To remove the dust from a dusty image and make the scene clear, this paper proposes an effective dedusting network for single image (SIDNet). The SIDNet develops several dust-aware representation extraction (DustAre) modules with the same structure. Each DustAre module contains two branches. The first branch estimates global veiling-light and local spatial information. The second branch generates a dust-aware map, and fuses the global veiling-light and the local spatial information to generate the output. Moreover, the SIDNet introduces the color cast correction scheme to the neural network for further improving the image dedusting performance on real dusty images. Under the consideration of the differences of a^* and b^* components between a dusty image and a dust-free image in CIELAB color space, the SIDNet designs a new loss function to better guide network training. In addition, after considering the relationship of the scene depth between a real dusty image and a dust-free image, we developed a novel dusty image synthesis method for enriching dusty image datasets. Experimental results on synthetic dusty images and real dusty images show that the SIDNet dramatically improves dedusting performance than the-state-of-art methods. In the future, we will further explore semi-supervised deep learning networks for dusty image synthesis and image dedusting.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Jiayan Huang: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing. **Haiping Xu:** Data curation, Visualization, Supervision. **Guanghai Liu:** Validation, Supervision. **Chuansheng Wang:** Formal analysis, Investigation, Data curation, Validation, Supervision. **Zhongyi Hu:** Visualization, Supervision, Funding acquisition. **Zuoyong Li:** Resources, Supervision, Funding acquisition.

Acknowledgments

This work is partially supported by National Natural Science Foundation of China (61972187, 61866005), Natural Science Foundation of Fujian Province (2020J02024 and 2019J01756), Fuzhou Science and Technology Project (2020-RC-186), and Key Project of Zhejiang Provincial Natural Science Foundation (LD21F020001).

References

- [1] G. Liu, J. Yang, Deep-seated features histogram: a novel image retrieval method, *Pattern Recognit.* 116 (2021) 107926.
- [2] Y. Zheng, J. Fan, J. Zhang, X.-B. Gao, Exploiting related and unrelated tasks for hierarchical metric learning and image classification, *IEEE Trans. Image Process.* 29 (1) (2020) 883–896.
- [3] G. Liu, J. Yang, Exploiting color volume and color difference for salient region detection, *IEEE Trans. Image Process.* 28 (1) (2019) 6–16.
- [4] J. Wang, Y. Pang, Y. He, C. Liu, Enhancement for dust-sand storm images, *Int. Conf. Multimedia Model.* (2016) 842–849.
- [5] G. Gao, H. Lai, Z. Jia, Y. Liu, Y. Wang, Sand-dust image restoration based on reversing the blue channel prior, *IEEE Photonics J.* 12 (2) (2020) 1–16.
- [6] C. Wang, Z. Li, J. Wu, H. Fan, G. Xiao, H. Zhang, Deep residual haze network for image dehazing and deraining, *IEEE Access* 8 (2020) 9488–9500.
- [7] B. Li, X. Peng, Z. Wang, J. Xu, D. Feng, AOD-Net: all-in-one dehazing network, *IEEE Int. Conf. Comput. Vision* (2017) 4770–4778.
- [8] X. Qin, Z. Wang, Y. Bai, X. Xie, H. Jia, FFA-Net: feature fusion attention network for single image dehazing, *AAAI Conf. Artif. Intell.* 34 (7) (2020) 11908–11915.
- [9] X. Fu, B. Liang, Y. Huang, X. Ding, J. Paisley, Lightweight pyramid networks for image deraining, *IEEE Trans. Neural Netw. Learn. Syst.* 31 (6) (2019) 1794–1807.
- [10] Z. Fan, H. Wu, X. Fu, Y. Huang, X. Ding, Residual-guide network for single image deraining, *26th ACM Int. Conf. Multimedia* (2018) 1751–1759.
- [11] D. Berman, D. Levy, S. Avidan, T. Treibitz, Underwater single image color restoration using haze-lines and a new quantitative dataset, *IEEE Trans. Pattern Anal. Mach. Intell.* 43 (8) (2021) 2822–2837.
- [12] J. Huang, Z. Li, C. Wang, Z. Yu, X. Cao, FFNet: a simple image dedusting network with feature fusion, *Concurrency Comput.* (2021) e6462.
- [13] K.M. Jeong, B.C. Song, Fog detection and fog synthesis for effective quantitative evaluation of fog-detection-and-removal algorithms, *IEIE Trans. Smart Process. Comput.* 7 (5) (2018) 350–360.
- [14] F. Gasparini, R. Schettini, Color correction for digital photographs, *12th Int. Conf. Image Anal. Process.* (2003) 646–651.
- [15] K. Iqbal, M. Odetayo, A. James, R.A. Salam, A.Z.H. Talib, Enhancing the low quality images using unsupervised colour correction method, *IEEE Int. Conf. Syst. Man Cybern.* (2010) 1703–1709.
- [16] Q. Deng, Z. Huang, C.-C. Tsai, C.-W. Lin, HardGAN: a haze-aware representation distillation GAN for single image dehazing, *Eur. Conf. Comput. Vision* (2020) 722–738.
- [17] H. Koschmieder, Theorie der horizontalen sichtweite, *beitrage zur physik der freien atmosphare*, *Meteorol. Z.* 12 (1924) 3353.
- [18] R. Fattal, Single image dehazing, *ACM Trans. Graph. (TOG)* 27 (3) (2008) 1–9.
- [19] C. Godard, O. Mac Aodha, M. Firman, G.J. Brostow, Digging into self-supervised monocular depth estimation, *IEEE/CVF Int. Conf. Comput. Vision* (2019) 3828–3838.
- [20] W. Wang, X. Yue, H. Liu, Z. Pan, D. Tang, Y. Wang, D.U. Raoguo, S.U. Hongmei, F. Qian, S. Kazuhiko, Study on pollution characteristics of aerosols during sand-dust storm weather in beijing, *Acta Scientiae Circumstantiae* 22 (4) (2002) 494–498.
- [21] B. Cai, X. Xu, K. Jia, C. Qing, D. Tao, DehazeNet: an end-to-end system for single image haze removal, *IEEE Trans. Image Process.* 25 (11) (2016) 5187–5198.
- [22] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, B. Schiele, The cityscapes dataset for semantic urban scene understanding, *IEEE Conf. Comput. Vision Pattern Recognit.* (2016) 3213–3223.
- [23] N. Silberman, D. Hoiem, P. Kohli, R. Fergus, Indoor segmentation and support inference from RGBD images, *Eur. Conf. Comput. Vision* (2012) 746–760.
- [24] R. Girshick, Fast R-CNN, *IEEE Int. Conf. Comput. Vision* (2015) 1440–1448.
- [25] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., ImageNet large scale visual recognition challenge, *Int. J. Comput. Vis.* 115 (3) (2015) 211–252.
- [26] H. Zhang, V.M. Patel, Density-aware single image de-raining using a multi-stream dense network, *IEEE/CVF Conf. Comput. Vision Pattern Recognit.* (2018) 695–704.
- [27] R. Li, S. You, X. Zhang, Y. Li, Learning to dehaze from realistic scene with a fast physics-based dehazing network, *arXiv preprint arXiv:2004.08554* (2020).
- [28] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Image Process.* 13 (4) (2004) 600–612.
- [29] L. Liu, B. Liu, H. Huang, A.C. Bovik, No-reference image quality assessment based on spatial and spectral entropies, *Signal Process. Image Commun.* 29 (8) (2014) 856–863.
- [30] C.E. Shannon, A mathematical theory of communication, *ACM SIGMOBILE Mob. Comput. Commun. Rev.* 5 (1) (2001) 3–55.
- [31] D.G. Lowe, Distinctive image features from scale-invariant keypoints, *Int. J. Comput. Vis.* 60 (2) (2004) 91–110.
- [32] S. Zhao, L. Zhang, Y. Shen, Y. Zhou, RefineDNet: a weakly supervised refinement framework for single image dehazing, *IEEE Trans. Image Process.* 30 (2021) 3391–3430.