ORIGINAL PAPER

DCNet: Dark Channel Network for single-image dehazing

Akshay Bhola[1] [ORCID] · **Teena Sharma[2]** · **Nishchal K. Verma[2]**

Abstract

Single-image dehazing is an extensively studied field and an ill-posed problem faced by vision-based systems in an outdoor environment. This paper proposes a dark channel network to estimate the transmission map of an input hazy scene for single-image dehazing. The architecture constitutes two major components—feature extraction layer and convolutional neural network layer. The former extracts the haze relevant features, while latter convolve these features with filter kernels to estimate the true scene transmission. Finally, the estimated transmission map is used to obtain the dehazed image using atmospheric scattering model. The experiments have been performed on synthetic hazy images and benchmark hazy dataset available in the literature. The performance of the proposed architecture outperforms the existing models in terms of standard quantitative metrics—mean square error, structural similarity index, and peak signal-to-noise ratio.

Keywords Dark channel prior · Single-image dehazing · CNN · Dehazing · Image restoration

1 Introduction

The performance of computer vision-based systems is sensitive to the visual quality of the observed scene for various inference-related tasks such as object detection and image segmentation. Poor visibility of observed scene degrades the performance of algorithms developed for these applications. Nowadays, vision-based systems rely on scene enhancement techniques to rectify the captured noisy observations and restore the clear scene. Physically, noise is induced by the suspended atmospheric particles such as dust and smoke. The size, material, shape, and concentration of these particles affect the intensity of the added noise [1]. During image acquisition, light rays reflected from object surface suffer scattering and absorption due to these suspended particles. It results in poor contrast and visibility in the acquired image.

This phenomenon of dispersion of light due to suspended atmospheric particles is called atmospheric scattering. Mathematically, the atmospheric scattering phenomenon has been formulated by Narasimhan and Nayar [2,3] as:

$$S^\lambda(a, b) = J^\lambda(a, b)t(a, b) + A(1 - t(a, b)) \tag{1}$$

where $S^\lambda(a, b)$ is the observed scene, $J^\lambda(a, b)$ is the true scene radiance, A is the global atmospheric airlight, $\lambda \in$ {R,G,B} represent the associated color channel, (a, b) is a pixel location, and $t(a, b)$ is the scene transmission defined as:

$$t(a, b) = \exp(-\beta \times d(a, b)) \tag{2}$$

where β is the scattering coefficient and $d(a, b)$ is the scene depth at pixel location (a, b).

1.1 Literature review

Based on the atmospheric scattering model, researchers came up with various dehazing methods categorized as image enhancement-based, priors or assumption-based, fusion-based, and deep learning-based methods. Before Narasimhan and Nayar's haze formation model, Kopeika [4] and Yitzhaky [5] used weather predicted atmospheric modulation transfer function and prior distance estimate. Further, Oakley et al. [6]

✉ Akshay Bhola
akbhola.bhola@gmail.com

Teena Sharma
teenashr@iitk.ac.in

Nishchal K. Verma
nishchal@iitk.ac.in

1   Department of Computer Science and Mechanical
     Engineering, Indian Institute of Technology Kanpur, Kanpur,
     India

2   Department of Electrical Engineering, Indian Institute of
     Technology Kanpur, Kanpur, India

and Tan et al. [7,8] developed a physics-based model to enhance the hazy scene without any weather prediction. Earlier approaches for image dehazing require additional information such as Schechner and Shwartz [9,10] use multiple images of a scene with various degrees of polarization using polarized filters to obtain the clear image. In real-time applications, vision-based algorithms cannot rely on these types of algorithms. Hence, it generates the need for single-image dehazing.

Single-image dehazing is an ill-posed problem in which relevant feature(s) of an image is modeled based on the available pixel information, which is utilized to obtain a clear image from an input hazy image. Different researchers focus on different properties of image features and utilizes different modeling techniques to target this problem. One of the associated effect of haze is the loss of contrast in the acquired images. Based on this, Tan [11] proposed a single-image dehazing framework by maximizing the local contrast of an image, and Fattal [12] proposed to enhance the scene visibility by estimating its true transmission. Alternatively, for the estimation of true transmission, He et al. [13] introduced dark channel prior (DCP)-based method for single-image dehazing. The performance of DCP was further improved by bilateral [14], median [15,16], edge-preserving [17], and guided [18] filters. A fast hazy region identification technique based on semi-inverse of an image was proposed by Ancuti et al. [19]. Further, they came up with a multi-scale fusion-based approach to restore the clear image [20]. The dehazing quality was also improved by introducing a boundary constraint and contextual regularization (BCCR) [21]. Since all of these methods depend on various priors and assumptions, they cannot be applied globally. Tang et al. [22] trained a random forest-based regression model using various haze relevant features for estimating true scene transmission. This learning-based framework became the motivation for the use of deep learning-based architectures for image dehazing. Cai et al. [23] proposed DehazeNet architecture for estimating true scene transmission using convolutional neural network (CNN)-based architecture. Similarly, Ren et al. [24] developed an end-to-end multi-scale CNN for estimating scene transmission. Following the success of CNN-based models for single-image dehazing task, a number of different novel architectures have been proposed. AIPNet [25] is based on the observation that the illumination channel gets significantly altered by haze compared to any other channels in YCrCb color space. Li et al. [26] modeled a novel variable instead of transmission map and proposed a fast and effective CNN architecture to model it. GCANet [27] and GMAN [28] are based on encoder–decoder architecture to extract the noise component and to rectify it from the image.

1.2 Motivation

As mentioned earlier, single-image dehazing is an ill-posed problem with clear scene radiance, atmospheric light intensity and transmission map being the unknown entities. Looking at (1), clear scene $J^\lambda(a, b)$ depends on the observed scene $S^\lambda(a, b)$, transmission $t(a, b)$, and global atmospheric airlight A as:

$$J^\lambda(a, b) = \frac{S^\lambda(a, b) - A\left(1 - t(a, b)\right)}{t(a, b)}. \tag{3}$$

The interpretation behind (1) is that the original clear scene radiance $J^\lambda(a, b)$ gets degraded by the scattering caused by suspended particles in the air. Additionally, these suspended particles scatter the atmospheric light, which causes diffused white light in the observed scene. In (3), transmission map represents a crucial unknown entity required to retrieve the clear scene. Therefore, modeling of transmission map is carried out in the following study. Atmospheric light varies minimally across the scene and thus approximated as a constant.

1.3 Contributions

The main objective of the vision-based algorithms is to perceive the visual stimuli from the environment in a similar way as humans perceive. The human brain can easily identify hazy regions and is able to distinguish varying haze densities. The CNN-based models utilize the local information of an image in a similar manner as that of a human eye. Therefore, this paper proposes a CNN-based architecture for the estimation of true scene transmission. The architecture is named as "dark channel network (DCNet)". Figure 1 shows the diagram of the proposed model. The contributions of this paper are summarized as:

– This paper proposes a novel CNN-based architecture named as "dark channel network (DCNet)" for the estimation of true scene transmission of a hazy scene with minimum number of layers in comparison with the CNN-based dehazing models available in the literature. The proposed DCNet architecture comprises of two layers: (1) feature extraction layer and (2) convolutional neural network layer.
– The proposed network has been extensively analyzed, and performance is compared with current state-of-the-art models over the standard benchmark datasets to show the effectiveness of the model and to verify the proposed hypothesis.

The paper is organized as follows in further sections: Section 2 explains the architecture of the proposed DCNet.
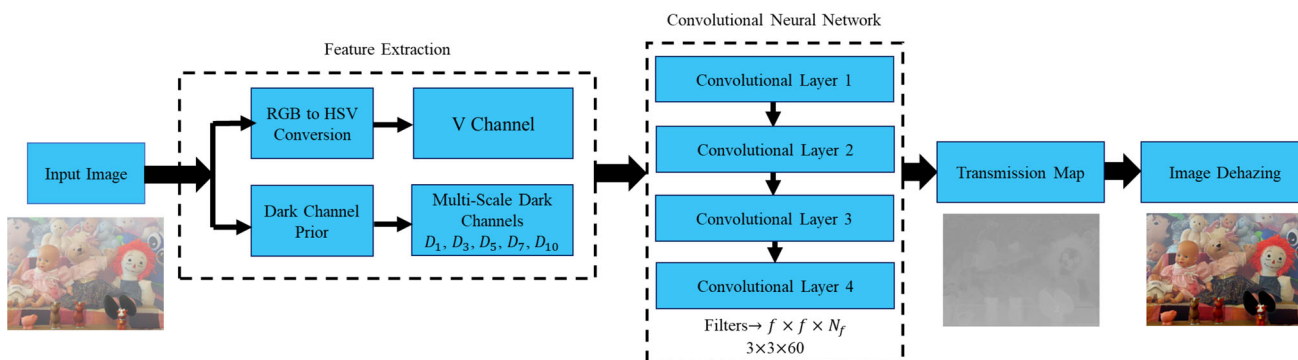
Fig. 1 Block diagram of the proposed DCNet. D_p is referred as dark channel prior for patch size $p = \sqrt{p} \times \sqrt{p}$ where $p = 1 \times 1, 3 \times 3, 5 \times 5, 7 \times 7, 10 \times 10$

Sections 3 and 4 discuss the training and experimental results of the model. Section 5 concludes the paper with its limitations and future scope.

2 DCNet: proposed architecture

The DCNet architecture comprises of two layers:

1. Feature extraction layer
2. Convolutional neural network layer

In the following subsections, these layers are discussed in detail.

2.1 Feature extraction layer

Feature extraction layer extracts relevant pixel maps for rectifying haze from an input hazy image. These features constitutes value (V) channel and multi-scale dark channels.

2.1.1 Value (V) channel

In general, the V channel or illumination channel of an image constitutes edge information for textural interpretation. In HSV color domain, V channel of an image constitutes maximum edge information (Refer Fig. 2). Mathematically, the V or illumination channel is defined as:

$$S^V(a, b) = \max_{\lambda \in \{R,G,B\}} S^\lambda(a, b). \tag{4}$$

The observed scene in hazy environment involves degraded radiance of scene due to scattering of light, as discussed in the previous section. It causes loss of contrast and saturation, along with significant variation in value and saturation channels of an image. The mean square error (MSE) plot for 100 images from NYU dataset [30] is shown in Fig. 3. The
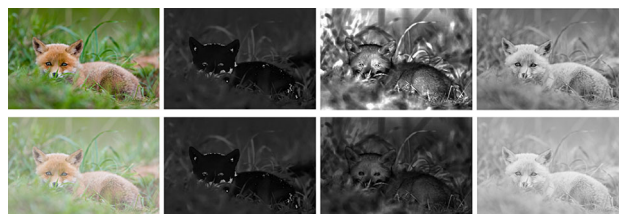

Fig. 2 Visualization of an image in RGB and HSV color domains. Top row: clear scene image. Bottom row: hazy scene image. From left to right: RGB image, H channel, S channel, and V channel
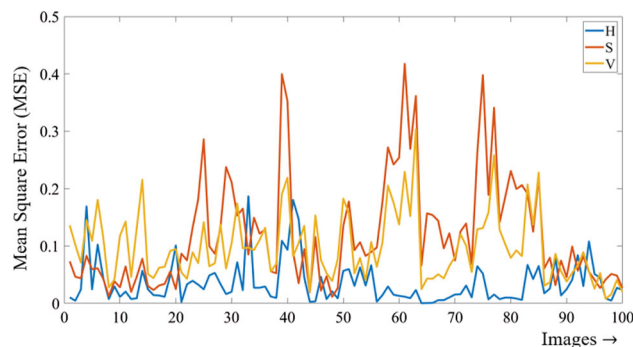

Fig. 3 Plots for the mean square error obtained between clear and hay images for H, S, and V channels

MSE is evaluated for H, S, and V channels of clear and hazy images. Clearly, the V channel is significantly affected by the presence of haze. Another critical property of V channel is that it provides a means to identify originally bright regions in the image and to distinguish it from hazy regions.

The minimum channel for the atmospheric scattering model (1) is defined as:

$$S_{\min}(a, b) = J_{\min}(a, b)t(a, b) + A(1 - t(a, b)) \tag{5}$$

where

$$J_{\min}(a, b) = \min_{\lambda \in \{R,G,B\}} S^\lambda(a, b). \tag{6}$$

The maximum channel for the atmospheric scattering model (1) is defined as:

$$S_{max}(a, b) = J_{max}(a, b)t(a, b) + A(1 - t(a, b)) \qquad (7)$$

where

$$J_{max}(a, b) = \max_{\lambda \in \{R,G,B\}} S^{\lambda}(a, b). \qquad (8)$$

The difference between (5) and (7) is

$$\Delta S(a, b) = S_{max}(a, b) - S_{min}(a, b) \qquad (9)$$

or

$$\Delta S(a, b) = (J_{max}(a, b) - J_{min}(a, b)) t(a, b). \qquad (10)$$

It is to be noted that $\Delta S(a, b)$ have smaller magnitude for bright regions and significantly larger magnitude for dark regions. Clearly, the value channel comprises relevant image information for image dehazing. In the proposed architecture, RGB space is transformed to HSV space and the value channel is extracted by

$$\left(S^R, S^G, S^B\right) \to \left(S^H, S^S, S^V\right). \qquad (11)$$

2.1.2 Multi-scale dark channels

Dark channel of an image is the pixel-wise minima over a local patch at corresponding pixel location of RGB channels [13]. It is defined as:

$$D_p(a, b) = \min_{(i, j) \in \Omega_p(a,b)} \left(\min_{\lambda \in \{R,G,B\}} \left(S^{\lambda}(i, j) \right) \right) \qquad (12)$$

where $S^{\lambda}(i, j)$ represents the observed image intensity at location (i, j), $\lambda \in \{R,G,B\}$, and $\Omega_p(a, b)$ is a patch of size $\sqrt{p} \times \sqrt{p}$ centered at (a, b) in $S^{\lambda}(i, j)$.

The significance of dark channel is that it provides a way of capturing haze density in a hazy image. For a clear image, the image intensity over a local path at each pixel location, except the sky, is close to zero for at least one of the color channels (RGB) [13]. In a hazy environment, the scattered light intensity adds up to the RGB channels of an observed scene. Thus, the pixel intensity of all 3 color channels gets incremented corresponding to the depth of the scene [13,22]. Figure 4 shows the dark channels of an image for different window sizes (p). He et al. [13] stated the following

– For smaller p, dark channel property fails in certain regions, but the pixel map constitutes better edge information.
– For larger p, dark channel property holds better, but dehazed image shows "halo effects".

Fig. 4 Visualization of dark channel $D_p(a, b)$ obtained for different patch sizes. From left to right and top to bottom: hazy scene image, $D_p(a, b)$ for $p = \sqrt{p} \times \sqrt{p} = 1 \times 1, 3 \times 3, 5 \times 5, 7 \times 7, 9 \times 9, 11 \times 11, 13 \times 13, 15 \times 15, 17 \times 17, 19 \times 19$, and 21×21

The feature extraction layer of DCNet extracts multi-scale dark channels. It takes advantage of both small patch size and large patch size to extract out the information from input hazy scene, thus resulting in a transmission map with sharp edges. The proposed DCNet has been analyzed over various combinations of multi-scale dark channels. It has been observed that the network performance improves by increasing the number of dark channels. The increase tends to saturate after the patch size of $p = 10 \times 10$. In this paper, DCNet is designed by choosing dark channels with patch sizes $p \in \{1 \times 1, 3 \times 3, 5 \times 5, 7 \times 7, 10 \times 10\}$.

The output of feature extraction layer comprises the concatenation of the V channel and multi-scale dark channels of an input hazy image. It can be represented as:

$$F(a, b, z) \leftarrow \left(S^V(a, b), D(a, b)\right) \qquad (13)$$

where $D(a, b)$ represent the set of all dark channels and z represents the number of layers being concatenated (i.e., $z = 6$ in the proposed architecture).

2.2 Convolutional neural network layer (CNN layer)

The CNN layer of DCNet architecture constitutes the primary component involved in estimating the transmission map. The output of feature extraction layer acts as the input to CNN layer. The operation of a convolution layer is defined as:

$$C_l = f(W_l * C_{l-1} + b_l) \qquad (14)$$

where W_l and b_l represent the weights and biases, respectively, of the lth convolutional layer, f represents the activation function, and C_l represents the output of lth convolutional layer. The output of the first convolutional layer is

$$C_1 = f(W_1 * C_0 + b_1) \qquad (15)$$

Table 1 Architecture of the proposed model

Layer	Filter size
1st Layer	$64 \times 3 \times 3$
2nd Layer	$32 \times 3 \times 3$
3rd Layer	$16 \times 3 \times 3$
4th Layer	$8 \times 3 \times 3$
5th Layer	$1 \times 3 \times 3$

where

$$C_0 = F(a, b, z). \tag{16}$$

Convolutional network constitutes 5 convolutional layers with filter size of $f \times f = 3 \times 3$, as shown in Table 1. Zero padding is applied to maintain the output dimensions across layers. The output of CNN layer is the estimated transmission map of the image as:

$$t = C_4 = f(W_4 * C_3 + b_4). \tag{17}$$

Finally, the dehazed image is obtained using (3), with global atmospheric airlight as $A = (0.95, 0.95, 0.95)$. The variation of atmospheric light is usually minimal, thus kept constant in the model.

3 Training of DCNet

The training of DCNet has been carried out as a supervised learning problem. The model is trained over synthetically generated hazy patches of images with their corresponding transmission maps. These patches are extracted by randomly cropping small sections of images of size 20×20 and then adding uniform haze in the patches using (1) for a random value of transmission map t. Figure 5 shows some sample patches. The motivation behind this formulation is that usually the variation of depth over a small patch of an image tends to be insignificant. Thus, the transmission map over the region is assumed to be a constant.

The training of the model has been carried out using two different specifications for drawing random transmission maps. The primary model is trained by drawing transmission map uniformly randomly such that $t \in [0.4, 1.0]$. This model is more robust and better handles heavy haze situations. The secondary model is trained by drawing $t \in [0.7, 1.0]$, thus better suited for dehazing of lightly hazed images.

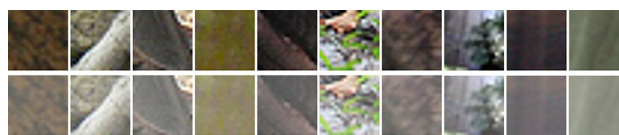In further sections, the dataset and model analysis have been discussed in detail.

Fig. 5 Visualization of training patches of size 20×20. Top row: clear patches. Bottom row: corresponding hazy patches

3.1 Dataset, loss function and optimizer for DCNet

The training dataset is created by custom selection of a subset of relevant images from "ImageNet 2013 validation set" [31]. The dataset comprises of 200 images, and in total 10,000 random patches, i.e., 50 patches per image, are extracted for the training of DCNet. Similarly, the Synthetic Test dataset is created by custom selection of 250 distinct and challenging images from "ImageNet 2013 validation set".

The training of the proposed model is carried out using the MSE loss function (18). The MSE loss function is defined as:

$$L(\Theta) = \frac{1}{2} \sum_{n=1}^{N} ||NN(x_n) - t_n||^2. \tag{18}$$

The dataset is randomly shuffled over each epoch, and RMSprop optimizer in PyTorch library is utilized during the training procedure.

3.2 Architecture analysis

The proposed DCNet has various hyper-parameters which affects the effectiveness and speed of the proposed model significantly. The network has been analyzed over these hyper-parameters to obtain the final model architecture. These parameters are discussed below.

3.2.1 Haze relevant features

The model has been analyzed over a number of distinct features, i.e., value (V) channel, multi-scale dark channels, and multi-scale saturation (S) channels. Analysis shows that adding multi-scale S channels as another feature to the CNN layer, in addition to the V and multi-scale dark channels, results in negligible improvement in the performance of the model. Thus, the final architecture of the proposed model constitutes only the value channel and multi-scale dark channels.

3.2.2 Patch size of dark channel

The patch size of dark channel is an important parameter of the feature, and the behavior of dark channel is highly sensitive to it. The small-sized patch constitutes edge information

Table 2 Performance of DCNet for different combinations of dark channels (D_p) in DCNet

No. of dark channels	Patch size $p = \sqrt{p} \times \sqrt{p}$	MSE
4	$\sqrt{p} \in \{1, 3, 5, 7\}$	0.00716
5	$\sqrt{p} \in \{1, 3, 5, 7, 10\}$	**0.00198**
6	$\sqrt{p} \in \{1, 3, 5, 7, 10, 15\}$	0.00189

Bold represents performance of the proposed architecture

but the dark channel hypothesis fails under this specifications, whereas large-sized patch satisfies the dark channel hypothesis but loses the boundary information. To utilize the benefit of both the features while mitigating their limitations, a hybrid implementation of multi-scale dark channels has been proposed. The study has been conducted to analyze the performance of the proposed model over varying set of dark channels ($D_t | t \in \mathbf{S}$), where $\mathbf{S}$ represents a set of window sizes. The study shows that the performance of the model improves with increasing number of dark channels. The improvement becomes minimal after $p = 10 \times 10$, whereas the computational load keeps increasing with increase in the number of dark channels. Table 2 shows the performance of DCNet for various combinations of dark channels. The model is optimized over accuracy and computational load. The final architecture constitutes a set of 5 dark channels, i.e., ($D_1, D_3, D_5, D_7, D_{10}$).

3.2.3 Number of convolutional layers

It is a general trend in neural networks that with increasing complexity of the model, the performance of model improves. Concurrently, the computational load of network also increases. In this work, the proposed model has been analyzed for multiple sets of convolutional layers. The performance of the model shows an increasing trend, and it tends to saturate after $l = 5$ layers. Thus, in the final architecture, 5 convolutional layers are used.

3.2.4 Activation function

The proposed architecture has been analyzed with a number of different activation functions, i.e., ReLU, leaky ReLU, and sigmoid function. Figure 6 shows the behavior of these functions. Analysis shows that the proposed architecture with leaky ReLU function outperforms models with similar architecture constituting sigmoid and ReLU functions. The superior performance over ReLU is due to the small negative slope of leaky ReLU function which helps in bypassing the dead ReLU problem. The proposed model constitutes each convolutional layer being interleaved by a leaky ReLU function with a slope of 0.05.
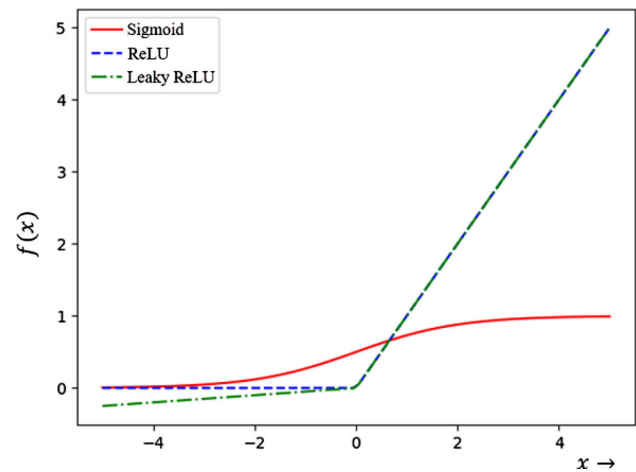

Fig. 6 Plots of sigmoid, ReLU, and leaky ReLU activation functions

4 Results and discussions

Extensive experiments have been conducted to comprehend the effectiveness of the proposed model. The performance of the proposed model has been analyzed and compared with the current state-of-the-art methods: DCP [13], BCCR [21], SIDMF [20], CAP [29], MSCNN [24], DehazeNet [23], GCANet [27], AOD-Net [26], and GMAN [28]. Experiments have been conducted on the Synthetic Test dataset [31], Middlebury Stereo dataset [32–34], SOTS and HSTS datasets [35] and some standard naturally hazed images. For fair comparison, the results over the state-of-the-art methods were reproduced with author's support. Both quantitative and qualitative analyses have been conducted to obtain the detailed comparison among these models. The quantitative analysis is conducted using standard statistical measures, i.e., mean square error (MSE), structural similarity index (SSIM) [36], and peak signal-to-noise ratio (PSNR).

Tables 3 and 4 show the average performance comparison of DCNet with current state-of-the-art methods over Synthetic Test dataset and Middlebury Stereo dataset, respectively. The hazy images are generated using (1) with $t = 0.6$. The results show that DCNet outperforms all other models.

Figure 7 shows the visual comparison of the effectiveness of models over some images from Middlebury Stereo dataset. Different models are able to remove haze to varying extents. DCP and BCCR models removed haze to some extent but affected the color composition significantly. SIDMF removed the haze non-uniformly and affected the other color channels. CAP, MSCNN, DehazeNet, GCANet, AOD-Net, and GMAN are able to remove significant haze, but the resulting images are degraded to some extent in their color composition and texture information. The resulting images generated using DCNet are fairly close to the original clear images preserving the texture, color, and edge informa-

Table 3 Average performance comparison of DCNet with state of the art on Synthetic dataset

Metric	DCP [13]	BCCR [21]	SIDMF [20]	CAP [29]	MSCNN [24]	DehazeNet [23]	GCANet [27]	AOD-Net [26]	GMAN [28]	DCNet
MSE	0.0075	0.0062	0.0242	0.0024	0.0054	0.0043	0.0072	0.0044	0.0039	**0.0013**
SSIM	0.8472	0.7708	0.3847	0.9011	0.8283	0.9412	0.8664	0.9025	0.9023	**0.9571**
PSNR	19.6324	18.0292	11.5781	22.2278	19.2127	24.7768	23.1009	24.7332	24.8459	**28.9276**

Bold represent best performance for a particular metric

Table 4 Average performance comparison of DCNet with state of the art on Middlebury Stereo dataset

Metric	DCP [13]	BCCR [21]	SIDMF [20]	CAP [29]	MSCNN [24]	DehazeNet [23]	GCANet [27]	AOD-Net [26]	GMAN [28]	DCNet
MSE	0.0098	0.0072	0.0309	0.0016	0.0030	0.0030	0.0103	0.0026	0.0039	**0.0007**
SSIM	0.8431	0.7810	0.4808	0.9367	0.8677	0.9062	0.8021	0.9419	0.9201	**0.9830**
PSNR	16.7103	17.8482	10.8317	24.2569	21.2718	22.1496	20.6335	27.1277	24.2999	**32.1301**

Bold represent best performance for a particular metric

Fig. 7 Comparative visualization of dehazing methods on Baby and Drumsticks images from Middlebury Stereo dataset

tion. Similarly, Fig. 8 shows the performance comparison of DehazeNet and DCNet on a sample image of "bird". The zoomed views highlight the feathers in original clear image, its corresponding hazy image, DehazeNet output, and DCNet output. The output obtained using DehazeNet seems over-saturated, but DCNet output highly resembles the original clear image.
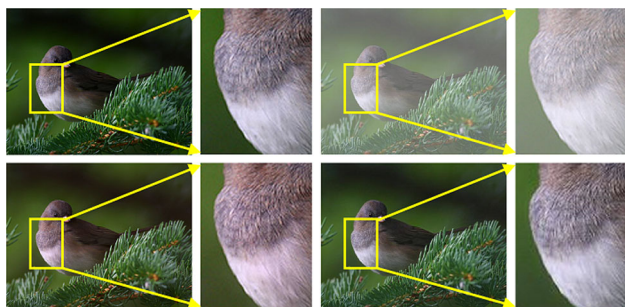
Fig. 8 Comparison of DehazeNet and DCNet on "bird" image. From left to right and top to bottom: Clear image and its feather region, hazy image and its feather region, dehazed image using DehazeNet and its feather region, and dehazed image using DCNet and its feather region

Table 5 Performance evaluation of DCNet on SOTS Outdoor Dataset

Metric	$\beta \in [0.6, 0.9]$	$\beta \in [1.0, 1.4]$	$\beta \in [1.5, 2.0]$
PSNR	28.74	24.42	21.98
SSIM	0.8871	0.8572	0.7926

Furthermore, the performance of DCNet has been analyzed on SOTS, HSTS, and standard benchmark datasets [35]. The PSNR and SSIM metrics of the model on SOTS outdoor dataset are provided in Table 5. Table 6 shows the results on Synthetic HSTS dataset. (The performance metrics of other state-of-the-art models are retrieved from [35].) The DehazeNet model outperforms all other models on HSTS dataset. It is to be noted that for consistency in experimental results throughout the paper, the evaluation is carried out using originally trained DCNet model (with synthetic dataset). Hence, this could be a reason for disparity in results.

Additionally, the proposed model has been analyzed on some standard natural images frequently used in the literature for comparison. These are interesting cases containing regions like clouds, marble, etc., which generally becomes the bottleneck for dehazing models in removing haze, as white regions are potential areas for the presence of heavy haze. Their corresponding clear images are not available. Figures 9 and 10 show the results for "Canyon" and "Girls" image, respectively. DCP model rectified hazy regions but, on the other hand, increased the contrast in certain areas, such as sky region, unrealistically high. BCCR shows similar behavior as that of DCP. In SIDMF, the resulting image shows non-uniform haze removal. In CAP, the model removed significant haze but also caused loss in edge information and over-saturation. MSCNN preserved the structural integrity and naturalness of the image, but the image still contains haze to some extent. DehazeNet effectively removed haze while maintaining the structural integrity and naturalness, but the image shows halo effects in certain regions. It is distinctly visible in high-frequency regions of the image. GCANet and GMAN effectively removed haze from the image, but the color composition is not preserved. AOD-Net is able to remove haze and better preserve the color composition of the image. In comparison with all aforementioned methods, DCNet is able to preserve the edge information, structural integrity, and is effectively being able to remove haze from the image while maintaining its naturalness. However, there are slight halo effects in the regions with sharp color change.

Table 6 Performance evaluation of DCNet on HSTS dataset

Metric	DCP [13]	BCCR [21]	CAP [29]	DehazeNet [23]	MSCNN [24]	AOD-Net [26]	DCNet
PSNR	14.84	15.08	21.53	**24.48**	18.64	20.55	22.27
SSIM	0.7609	0.7382	0.8726	**0.9153**	0.8168	0.8973	0.8003

Bold represent best performance for a particular metric
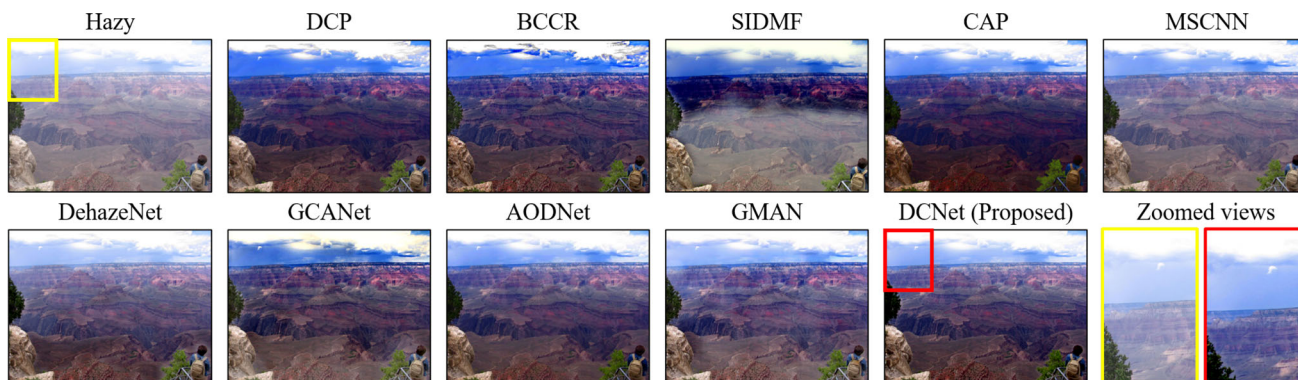

Fig. 9 Comparative visualization of dehazing methods on Canyon image. The zoomed views are shown for sky region in input hazy image (yellow box) and dehazed image using DCNet (red box) (color figure online)

Fig. 10 Comparative visualization of dehazing methods on Girls image. The zoomed views are shown for region with illumination and haze both, in input hazy image (yellow box) and dehazed image using DCNet (red box) (color figure online)

Fig. 11 Dehazing results on some standard images. Top row: hazy images; bottom row: dehazed images using DCNet

Table 7 Results of ablation studies on synthetic and Middlebury Stereo dataset

Metrics	Synthetic dataset		Middlebury dataset	
	DCNet†	DCNet	DCNet†	DCNet
MSE	0.0181	**0.0013**	0.0130	**0.0007**
SSIM	0.8211	**0.9571**	0.8790	**0.9830**
PSNR	21.53	**28.93**	21.92	**32.13**

Bold represent best performance for a particular metric
DCNet† DCNet without V channel

Figure 11 shows experimental results of DCNet on some other standard naturally hazed images. The resulting images from the model show that the closer regions have been successfully resolved, whereas in farther regions, light haze is still prevalent which maintains the naturalness in accordance with human visual perception.

4.1 Ablation studies

The proposed model takes 5 dark channels and the value channel as input features. To realize the importance of these features in the model, an ablation study is conducted. Table 7 shows the performance comparison of the proposed architecture with and without value channel as an input feature to the proposed network. It is quite evident that the performance in all three metrics significantly degrades without the value channel. Similarly, the analysis has been conducted to understand the importance of dark channels with different window sizes. Table 2 shows the performance of model with different set of dark channels. As stated earlier, the performance of the proposed model improves by adding more dark channel. The improvement becomes minimal after the window size of 10×10, but the computational load keeps increasing. Thus, the final architecture constitutes a set of 5 dark channels.

5 Conclusions

In this paper, a novel single-image dehazing model— "DCNet" has been proposed, constituting two major components, namely feature extraction layer and CNN layer, for estimating true transmission map of a hazy scene. The model learns nonlinear mapping between input hazy image and its transmission map, used for dehazing of image. The experimental results show that the proposed model outperforms the current state-of-the-art methods for single-image dehazing task over wide range of datasets.

Still, there is room for improvement in the model performance. The model shows slight halo effects in high-frequency regions of the images involving heavy hazy. Also, the model shows non-ideal performance in some images involving sizable bright regions.

References

1. Narasimhan, S.G., Nayar, S.K.: Vision and the atmosphere. Int. J. Comput. Vis. **48**, 233–254 (2002)
2. Nayar, S.K., Narasimhan, S.G.: Vision in bad weather. In: Proceedings of the Seventh IEEE International Conference on Computer Vision (ICCV), pp. 820–827. Kerkyra, Greece (1999)
3. Narasimhan, S.G., Nayar, S.K.: Chromatic framework for vision in bad weather. In: Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2000), pp. 598–605. Hilton Head Island, SC, USA (2000)
4. Kopeika, N.S.: A System Engineering Approach to Imaging. SPIE Press, Bellingham (1998)

5. Yitzhaky, Y., Dror, I., Kopeika, N.S.: Restoration of atmospherically blurred images according to weather-predicted atmospheric modulation transfer function. Optical Engineering **36**, 3064–3072 (1998)

6. Oakley, J.P., Satherley, B.L.: Improving image quality in poor visibility conditions using a physical model for degradation. IEEE Trans. Image Process. **7**, 167–179 (1998)

7. Tan, K., Oakley, J.P.: Enhancement of color images in poor visibility conditions. In: Proceedings 2000 International Conference on Image Processing (ICIP), p. 2. Vancouver, BC, Canada (2000)

8. Tan, K., Oakley, J.P.: Physics based approach to color image enhancement in poor visibility conditions. J. Opt. Soc. Am. A **18**(10), 2460–2467 (2001)

9. Schechner, Y.Y., Narasimhan, S.G., Nayar, S.K.: Instant dehazing of images using polarization. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), p. 325. Kauai, HI, USA (2001)

10. Shwartz, S., Namer, E., Schechner, Y.Y.: Blind haze separation. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2006), vol. 2, pp. 1984–1991. New York, NY, USA (2006)

11. Tan, R.: Visibility in bad weather from a single image. In: 2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2008), pp. 1–8. Anchorage, AK, USA (2008)

12. Fattal, R.: Single image dehazing. In: Proceedings on ACM SIGGRAPH'08, vol. 27, no 3. Los Angeles, California (2008)

13. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. IEEE Trans. Pattern Anal. Mach. Intell. **33**(12), 2341–2353 (2011)

14. Yu, J., Li, D., Liao, Q.: Physics-based fast single image fog removal. Acta Automatica Sinica **37**(2), 143–149 (2011)

15. Gibson, K.B., Vo, D.T., Nguyen, T.Q.: An investigation of dehazing effects on image and video coding. IEEE Trans. Image Process. **21**(2), 662–673 (2012)

16. Ding, M., Tong, R.: Efficient dark channel based image dehazing using quadtrees. Sci. China Inf. Sci. **56**(9), 1–9 (2013)

17. Shiau, Y.H., Yang, H.Y., Chen, P.Y., Chuang, Y.Z.: Hardware implementation of a fast and efficient haze removal method. IEEE Trans. Circuits Syst. Video Technol. **23**(8), 1369–1374 (2013)

18. He, K., Sun, J., Tang, X.: Guided image filtering. IEEE Trans. Pattern Anal. Mach. Intell. **35**(6), 1397–1409 (2013)

19. Ancuti, C.O., Ancuti, C., Hermans, C., Bekaert, P.: A fast semi-inverse approach to detect and remove the haze from a single image. In: Asian Conference on Computer Vision (ACCV 2010), Lecture Notes in Computer Science, vol. 6493, pp. 501–514. Springer, Berlin (2011)

20. Ancuti, C.O., Ancuti, C.: Single image dehazing by multiscale fusion. IEEE Trans. Image Process. **22**(8), 3271–3282 (2013)

21. Meng, G., Wang, Y., Duan, J, Xiang, S., Pan, C.: Efficient image dehazing with boundary constraint and contextual regularization. In: 2013 IEEE International Conference on Computer Vision (ICCV), pp. 617–624. Sydney, NSW, Australia (2013)

22. Tang, K., Yang, J., Wang, J.: Investigating haze-relevant features in a learning framework for image dehazing. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2014), pp. 2995–3002. Columbus, OH, USA (2014)

23. Cai, B., Xu, X., Jia, K., Qing, C., Tao, D.: DehazeNet: an end-to-end system for single image haze removal. IEEE Trans. Image Process. **25**(11), 5187–5198 (2016)

24. Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., Yang, M.-H.: Single image dehazing via multi-scale convolutional neural networks. In: European Conference on Computer Vision (ECCV), pp. 154–169 (2016)

25. Wang, A., Wang, W., Liu, J., Gu, N.: AIPNet: image-to-image single image dehazing with atmospheric illumination prior. IEEE Trans. Image Process. **28**(1), 381–393 (2019)

26. Li, B., Peng, X., Wang, Z., Xu, J., Feng, D.: AOD-Net: all-in-one dehazing network. In: 2017 IEEE International Conference on Computer Vision (ICCV), pp. 4780–4788. Venice, Italy (2017)

27. Chen, D., He, M., Fan, Q., Liao, J., Zhang, L., Hou, D., Yuan, L., Hua, G.: Gated context aggregation network for image dehazing and deraining. In: 2019 IEEE Winter Conference on Applications of Computer Vision (WACV). Waikoloa Village, HI, USA (2019)

28. Liu, Z., Xiao, B., Alrabeiah, M., Wang, K., Chen, J.: Single image dehazing with a generic model-agnostic convolutional neural network. IEEE Signal Process. Lett. **26**(6), 833–837 (2019)

29. Zhu, Q., Mai, J., Shao, L.: A fast single image haze removal algorithm using color attenuation prior. IEEE Trans. Image Process. **24**(11), 3522–3533 (2015)

30. Silberman, N., Hoiem, D., Kohli, P., Fergus, R.: Indoor segmentation and support inference from rgbd images, pp. 746–760. In: European Conference on Computer Vision. Springer, Berlin (2012)

31. Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L.: ImageNet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009). Miami, FL, USA (2009)

32. Scharstein, D., Szeliski, R.: High-accuracy stereo depth maps using structured light. In: 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2003), vol. 1, pp. 195–202. Madison, WI, USA, USA (2003)

33. Scharstein, D., Pal, C.: Learning conditional random fields for stereo. In: 2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2007). Minneapolis, MN, USA (2007)

34. Hirschmuller, H., Scharstein, D.: Evaluation of cost functions for stereo matching. In: 2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2007). Minneapolis, MN (2007)

35. Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W., Wang, Z.: Benchmarking single-image dehazing and beyond. IEEE Trans. Image Process. **28**(1), 492–505 (2019)

36. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)

Akshay Bhola received his B.Tech. degree with double major in Computer Science and Mechanical Engineering and minor in Cognitive Science from Indian Institute of Technology Kanpur, India. His research interests are in artificial intelligence, machine learning, deep learning, computer vision, natural language processing, and its applications.

Teena Sharma (Student Member, IEEE) received the B.Tech. degree in electronics and communication engineering from UPTU, Lucknow, India, in 2013 and the M.Tech. degree from Banasthali University, Vanasthali, India, in 2014. She is currently working toward the Ph.D. degree with the Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur, India. Her research interests are machine learning, computer vision, and its applications.

Nishchal K. Verma (Senior Member, IEEE) is currently a Professor with the Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur, India. He works in artificial intelligence (AI)-related theories and their practical applications to many inter-disciplinary domains like machine learning, deep learning, computer vision, prognosis and health management, bioinformatics, cyber-physical systems, complex and highly non-linear systems modeling, clustering, classifications, etc. He has authored or coauthored more than 230 research papers and four books (edited/coauthored) in the field of AI. Currently, he is serving as an Associate Editor for the IEEE Transactions on Artificial Intelligence, IEEE Transactions on Neural Networks and Learning Systems, and IEEE Computational Intelligence Magazine.