



CSIDNet: Compact single image dehazing network for outdoor scene enhancement

Teena Sharma¹  · Isha Agrawal² · Nishchal K. Verma¹

Received: 18 April 2020 / Revised: 19 July 2020 / Accepted: 29 July 2020 /
Published online: 17 August 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

This paper proposes a novel deep learning-based single image dehazing network named as Compact Single Image Dehazing Network (CSIDNet) for outdoor scene enhancement. CSIDNet directly outputs a haze-free image from the given hazy input. The remarkable features of CSIDNet are that it has been designed only with three convolutional layers and it requires lesser number of images for training without diminishing the performance in comparison to the other commonly observed deep learning-based dehazing models. The performance of CSIDNet has been analyzed on natural hazy scene images and REAListic Single Image DEhazing (RESIDE) dataset. RESIDE dataset consists of Outdoor Training Set (OTS), Synthetic Objective Testing Set (SOTS), and real-world synthetic hazy images from Hybrid Subjective Testing Set (HSTS). The performance metrics used for comparison are Peak Signal to Noise Ratio (PSNR) and Structural SIMilarity (SSIM) index. The experimental results obtained using CSIDNet outperform several well known state-of-the-art dehazing methods in terms of PSNR and SSIM on images of SOTS and HSTS from RESIDE dataset. Additionally, the visual comparison shows that the dehazed images obtained using CSIDNet are more appealing with better edge preservation. Since the proposed network requires minimal resources and is faster to train along with lesser run-time, it is more practical and feasible for real-time applications.

Keywords Outdoor scene enhancement · Dark channel · Illumination channel · Convolutional neural networks · Residual learning

✉ Teena Sharma
tee.shar6@gmail.com

Isha Agrawal
ishaagrwal2017@gmail.com

Nishchal K. Verma
nishchal@iitk.ac.in

¹ Department of Electrical Engineering IIT Kanpur, Kanpur, India, 208016

² Department of Computer Science and Engineering IIITDM Jabalpur, Jabalpur, India, 482005

1 Introduction

Various particulate matters such as dust, water drops, aerosol, etc. in the atmosphere often obscure the clarity of vision-based applications in outdoor environment. The most common phenomenon in outdoor environment due to inclement weather conditions is haze [25]. With a rise in the number of vision-based applications such as object classification [14], autonomous driving [31], remote sensing [5], etc., outdoor scene enhancement has become increasingly desirable for obtaining a clear scene. In literature, various methodologies have been developed to resolve the same.

1.1 Motivation

Haze is a signal dependent non-linear noise. This incurs attenuation in an image with increase in scene depth [25]. Thus, the pixel locations in a scene image suffer different amount of degradation. Single image dehazing has gained more popularity lately than those requiring additional data like multiple images [24] and different degrees of polarization [32]. Since acquiring additional information is not feasible for real-time applications, the process of single image dehazing becomes more challenging. A notable problem in dehazing is the absence of datasets with natural image pairs of hazy and haze-free images, as it is unlikely that the atmospheric conditions remain same on a hazy and a clear day. Consequently, synthetic hazy images are used for training purposes and then, the dehazing methods are tested on natural hazy scene images. Although significant work has been done to remove haze using deep learning, the difficulty still remains with the complicated structure of the architecture and requirement of rigorous training.

1.2 Contributions

This paper proposes a deep learning-based single image dehazing network named as “Compact Single Image Dehazing Network (CSIDNet)” for outdoor scene enhancement. The contributions of this paper are three-fold and summarized as follows:

- As the name implies, CSIDNet is a more compact network than the existing deep learning-based dehazing models and consists of only three convolutional layers.
- CSIDNet is trained on a much smaller dataset with lesser number of images without compromising the performance. Thus, it is easy to train with faster run-time and more approachable for real-time applications.
- The dehazed images obtained using CSIDNet are visually appealing and outperform the benchmarked deep learning-based dehazing models in terms of peak signal-to-noise ratio and structural similarity index measures.

Most of the dehazing models, to the best of our knowledge, are either computationally expensive, leading to an increased run-time, or require a number of resources for implementation. Contrary to this, CSIDNet has been designed with fewer layers and trained on an exceptionally lesser number of images, and yet significantly outperforms state-of-the-art methods.

The rest of the paper is organized as follows: Section 2 outlines the literature of image dehazing for outdoor scene enhancement, Section 3 presents the architectural design of proposed CSIDNet, Section 4 shows the comparison of results obtained with discussions, and finally, Section 5 highlights the concluding remarks with the future scope.

2 Related work

Some of the initial attempts for outdoor scene enhancement were based on Histogram Equalization (HE) and contrast restoration based methods [7, 26]. Tan [33] maximized the local contrast based on Markov random fields which led to over-saturated results. Fattal [6] used a refined image formation model to remove haze, but this is time consuming and fails in regions with dense haze. Meng et al. proposed Boundary Constraint and Contextual Regularization (BCCR) [23] method to efficiently remove haze with the assumption that haze-free images have better contrast than hazy ones. This result discontinuities in poor contrast regions. Ancuti et al. [2] proposed a contrast enhancement method to restore the discontinuities near edges lost due to poor contrast. Contrast restoration methods often produce unrealistic images due to the underlying assumption that the pixel intensity distribution of a clear scene must be uniform. The use of better assumptions and priors helped to make significant progress in outdoor scene enhancement. He et al. observed the low intensity values in the RGB image and proposed the Dark Channel Prior (DCP) [9]. DCP states that there always exist some pixels with low intensities within a local patch of one or more color channels in an RGB image. The drawback of DCP is haze overestimation in the sky regions. To reduce the computational time due to soft matting in DCP [15], He et al. introduced median of median filter [35], fast matting [8], and guided filter [10]. DCP was further used by Long et al. [19] to deal with halo artifacts by estimating an atmospheric veil for dehazing of remotely sensed hazy images. Thereafter, based on the behaviour of different image domains under hazy conditions, Tang et al. [34] proposed haze relevant features i.e. hue disparity, maximum saturation, and maximum contrast.

Recently, deep learning-based image dehazing models have achieved enormous popularity. Endeavours have been made to combine these models with the conventional atmospheric scattering model [25] for obtaining the clear scene. Zhu et al. introduced a Color Attenuation Prior (CAP) [38] based method to calculate the scene depth and then estimated the transmission map. However, it is not always accurate, and further calculation of airlight using this leads to accumulation and amplification of error. Cai et al. proposed DehazeNet architecture [4], which calculates the transmission map using four sequential operations. However, the dehazed images obtained using DehazeNet still persist some haze. Ren et al. proposed Multi-Scale Convolutional Neural Network (MSCNN) [29], which uses a combination of fine-scale and coarse-scale networks to output the clear scene. Li et al. combined the transmission map and atmospheric light as a new variable and used this to build an input adaptive model i.e. All-in-One Dehazing Network (AODNet) [16]. Ren et al. proposed Gated Fusion Network (GFN) [30], which is a supervised learning-based model taking three contrast relevant features [1, 28] as input to perform dehazing. Wang et al. introduced Atmospheric Illumination Prior Network (AIPNet) [36] based on the assumption that the luminance/illumination channel of a hazy image is much more affected by haze than its corresponding chrominance channel. Yang et al. [37] introduced a region detection network to approximate the transmission map, which was further used to enhance details in the dehazed image.

Yet, the relation of a hazy image and its corresponding haze-free image is quite complicated and difficult to interpret. This relation cannot be represented completely using the atmospheric scattering model [25] proposed by Narasimhan and Nayar for describing the haze formation phenomenon. As a result, the dehazing methods based on this model do not perform well on natural hazy images even if they show appreciable results on the synthetic images. In the recent literature, Liu et al. proposed Generic Model-Agnostic Network (GMAN) [18], which does not take any application specific features as input to restore the haze-free image. Since the performance of deep learning-based dehazing models depend

on the dataset of hazy and haze-free images, Li et al. proposed REalistic Single Image DEhazing (RESIDE) dataset [17]. The dataset consists of natural and synthetic hazy scene images of various haze levels with their ground truth clear counterparts. Recently, Qin et al. introduced Feature Fusion Attention Network (FFA-Net) [27] by combining the channel and pixel attentions for image restoration. Hence, the ultimate objective of outdoor scene enhancement is to increase the robustness of vision-based applications like object tracking. The enhancement+tracking pipeline aims to boost the real-time performance. Thus, the pipeline of faster enhancement and tracking results in the faster visual recognition. For example, the regression based networks for tracking with shrinkage loss [20] have gained attention among researchers. Furthermore, segmentation and tracking networks were proposed by Lu et al. [21, 22] in a unified and end-to-end trainable framework. Thus, the goal is to increase the performance of enhancement networks with faster run-time to synchronize with a faster vision-based applications.

3 Proposed network: compact single image dehazing network (CSIDNet)

This section includes explanation of the proposed dehazing network. Figure 1 outlines the architecture of the proposed network. In contrast to the other networks, CSIDNet comprises of only 3 convolutional layers. The detailed explanation of the architectural design of CSIDNet has been provided in the following subsections.

3.1 Hazy input features

Inspired by DCP [9] and AIPNet [36], CSIDNet extracts dark channel and illumination channel from input hazy image to learn the pattern of haze.

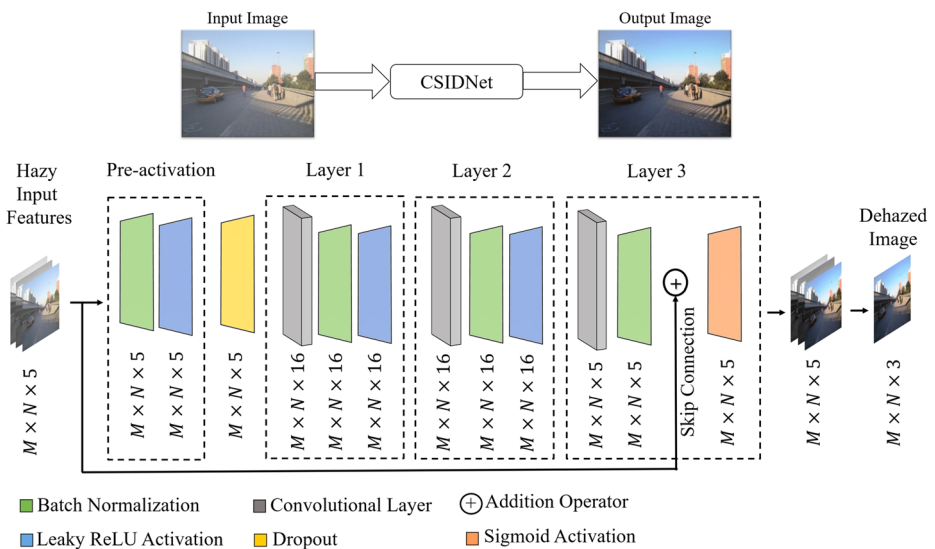


Fig. 1 Architecture of CSIDNet. The input image is of size $M \times N$ with five input features i.e. R, G, B channels of input hazy image (I), minimum channel (I_{\min}), and illumination channel (I_Y)

3.1.1 Dark channel prior

DCP [9] assumes that pixels of at least one color channel within a local patch always have low pixel intensity values in the non-sky region. Hence, the dark channel is obtained using

$$I_{\text{dark}}(a) = \min_{n \in P(a)} \left(\min_{C_h \in \{R, G, B\}} I^{C_h}(n) \right) \quad (1)$$

where, C_h represents R, G, and B color channels of the input hazy image $I(a)$, and $P(a)$ is a local patch centered at pixel location a . Equation (1) implies that the minimum intensity at each pixel location across all the color channels have a very low value in a haze-free image. This is mainly because of colourful objects dominated by one color channel and dark objects like tree trunks, shadows, etc. In this paper, the dark channel with patch size 1×1 has been considered.

3.1.2 Illumination channel

For a hazy input image $I(a)$, the illumination channel or Y channel is obtained by YCbCr color domain [3]. The RGB image channels are converted to YCbCr color channels using

$$\begin{bmatrix} I_Y(a) \\ I_{Cb}(a) \\ I_{Cr}(a) \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} I_R(a) \\ I_G(a) \\ I_B(a) \end{bmatrix} \quad (2)$$

where, a is the pixel location, $I_Y(a)$ is the illumination channel, $I_{Cb}(a)$ and $I_{Cr}(a)$ are the corresponding chrominance channels, and $I_R(a)$, $I_G(a)$, and $I_B(a)$ are the red, blue, and green color channels of the input image $I(a)$.

The RGB color channels, dark channel, and illumination channel are then concatenated to form the hazy input features for the network as

$$I_{\text{input}}(a) \xleftarrow{\text{Concatenate}} (I_R(a), I_G(a), I_B(a), I_{\text{dark}}(a), I_Y(a)). \quad (3)$$

3.2 Pre-activation

The hazy input features obtained by (3) are fully pre-activated [12] with batch normalization (Ψ) and leaky ReLU activation function (Φ) as

$$I_{\text{normalized}}(a) = \Phi(\Psi(I_{\text{input}}(a))). \quad (4)$$

A dropout layer (Ω) has also been included after the pre-activation stage to avoid any over fitting as

$$I_D(a) = \Omega(I_{\text{normalized}}(a)). \quad (5)$$

3.3 Convolutional layers

The output of dropout layer obtained by (5) is passed through three consecutive combinations of the following: convolutional layer, batch normalization layer, and activation layer (i.e. Layers 1, 2, and 3 of Fig. 1) as

$$I_{l+1}(a) = \Phi(\Psi(W_l * I_l(a))); \quad \forall l = 1, 2, 3 \quad (6)$$

where, l is the layer number, W is the kernel weight matrix between l and $l+1$ layer, $I_l(a)$ and $I_{l+1}(a)$ act as input and output of l^{th} layer, respectively. The input for the first layer is $I_1(a) = I_D(a)$ where, $I_D(a)$ is obtained from (5).

3.3.1 Skip connection

Furthermore, to compensate for any possible information loss, the network contains one global skip connection [13]. The output of this skip connection connects hazy input features $I_{\text{input}}(a)$ and output of 3rd convolutional layer after batch normalization as

$$I_{\text{skip}}(a) = \Psi(W_3 * I_3(a)) + I_{\text{input}}(a). \quad (7)$$

3.3.2 Sigmoid activation

The output of skip connection I_{skip} is then passed through Sigmoid activation (σ) to constraint the output in the range of 0 to 1 as

$$I_{\text{Sigmoid}}(a) = \sigma(I_{\text{skip}}(a)). \quad (8)$$

3.4 Dehazed image

The output $I_{\text{Sigmoid}}(a)$ contains 5 channels, where, the first three channels correspond to the required RGB dehazed output image as

$$I_{\text{dehazed}}(a) \xleftarrow{\text{Extract first three channels}} I_{\text{Sigmoid}}(a). \quad (9)$$

Thus, CSIDNet directly outputs a haze-free image $I_{\text{dehazed}}(a)$ in an interactive time without any intermediate results like transmission map or atmospheric light.

4 Results, validations, and discussions

This section presents qualitative and quantitative comparison of the results obtained using proposed network. The results are compared with the existing literature, namely BCCR [23], DCP [9], CAP [38], DehazeNet [4], MSCNN [29], AODNet [16], and GMAN [18].

While most of the deep learning-based dehazing models are trained on thousands to millions of images, CSIDNet has been trained only on 200 images without diminishing the performance. These training images have been selected randomly from Outdoor Training Set (OTS) of REalistic Single Image DEhazing (RESIDE) dataset [17]. The training of CSIDNet takes about 30 minutes on a system with Nvidia GeForce 940MX 2GB Graphics card. This is much lesser than the time taken by other models, which usually require at least 12-36 hours of training. The training has been conducted on an Intel Core i5-7200U 7th generation system with 2.5GHz processor and 8 GB DDR4 RAM. In further subsections, the explanation of datasets, network parameters, loss functions, quantitative and qualitative comparisons, and finally discussions have been provided.

4.1 Datasets

RESIDE dataset consists of 72,135 synthetic hazy images in Outdoor Training Set (OTS) for training. For testing, it contains 500 synthetic hazy images in Synthetic Objective Testing Set (SOTS) and 10 synthetic hazy images in Hybrid Subjective Testing Set (HSTS) with their respective ground truths.

For training of CSIDNet, 200 hazy images have been randomly selected from OTS to prepare one set of hazy images with their respective ground truths. Likewise, in total, five sets i.e. Set 1, Set 2, Set 3, Set 4, and Set 5 have been prepared randomly. The

Table 1 Performance comparison: Average PSNR AND SSIM measures (* First, † Second, and ††Third: Top 3 performance metrics values)

Methods	Testing Set			
	SOTS [17]		HSTS [17]	
	PSNR	SSIM	PSNR	SSIM
BCCR [23]	28.147	0.820	28.187	0.791
DCP [9]	16.620	0.818	14.840	0.761
CAP [38]	28.009	0.793	28.333	0.858 ^{††}
DehazeNet [4]	28.613 ^{††}	0.864 ^{††}	29.091*	0.890 [†]
MSCNN [29]	28.216	0.840	28.287	0.807
AODNet [16]	28.234	0.902*	28.254	0.891*
GMAN [18]	28.836 [†]	0.848	29.052 [†]	0.846
CSIDNet (Proposed)	28.995*	0.883 [†]	28.828 ^{††}	0.848

training sets are online available on the following link: <https://drive.google.com/open?id=1uCfliFpldUUWdzT5TKPX0kbVH5mMJ1iw>.

4.2 Network parameters and loss function

In CSIDNet, the depth of the convolutional layers are considered as 16 for layer 1, 16 for layer 2, and 5 for layer 3. The proposed network has been trained on images of size 224×224 , but can be tested for images of any resolution. The kernel size for the convolution has been chosen as 3×3 , the slope of Leaky ReLu is 0.2, and the dropout rate is 0.2. The kernel weights of convolutional layers have been initialized with He Uniform initializer [11]. CSIDNet is trained for 100 iterations with Adam optimizer of momentum values $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The loss function considered for the training of CSIDNet is Mean Square Error (MSE). It can be defined by

$$L_{MSE} = \frac{1}{C_h N_p} \sum_{i=1}^{C_h} \sum_{a=1}^{N_p} \left(I_{\text{dehazed}}^i(a) - I_{\text{gt}}^i(a) \right)^2 \tag{10}$$

where, I_{dehazed} is dehazed image, I_{gt} is ground truth, N_p represents the number of pixels in the image, and C_h denotes number of color channels. For further analysis, the proposed

Table 2 Performance comparison: Average run-time (in seconds)

Methods	Testing Set		
	SOTS [17]	HSTS [17]	Natural
BCCR [23]	1.174	1.002	1.604
DCP [9]	>50	>50	>50
CAP [38]	0.327	0.289	0.497
DehazeNet [4]	1.951	1.660	3.014
MSCNN [29]	0.954	0.742	1.689
AODNet [16]	0.349	0.297	0.562
GMAN [18]	1.441	1.373	2.036
CSIDNet (Proposed)	0.326	0.282	0.435

The bold entries are used to show better performance

network has also been trained with Mean Absolute Error (MAE) as a loss function. It is also termed as L_1 loss. It can be defined by

$$L_{L1} = \frac{1}{C_h N_p} \sum_{i=1}^{C_h} \sum_{a=1}^{N_p} \left| I_{\text{dehazed}}^i(a) - I_{\text{gt}}^i(a) \right| \quad (11)$$

The number of epochs used for training with L_1 loss function are 100, 150, and 200 in order to find the best possible hyper-parameters.

4.3 Quantitative comparison

The quantitative comparison of dehazed images obtained using CSIDNet has been analyzed in terms of Peak Signal to Noise Ratio (PSNR) and Structural SIMilarity (SSIM) index measures. The average PSNR and SSIM measures have been tabulated in Table 1 for images of SOTS and HSTS from RESIDE dataset [17]. Although the PSNR of GMAN is slightly



Fig. 2 Visual comparison on an image from SOTS of RESIDE dataset [17]

higher than the PSNR of CSIDNet for HSTS dataset, GMAN has a lower SSIM index. Lower SSIM accounts for the greater number of distortions. Similarly, the SSIM index of AODNet is higher than CSIDNet for SOTS dataset, but the run-time efficiency of CSIDNet is better than AODNet. Table 2 illustrates the comparison of average run-time on SOTS, HSTS, and natural hazy images.

A longer run-time implies that the network generates a lag in the process, leading to poor performance in real-time. Consequently, it is important for the output to be available in interactive time. As can be seen in Table 2, the run-time of DehazeNet is one of the highest amongst all the models, followed by GMAN, making it unfeasible for real-time purposes, whereas the proposed CSIDNet is fastest among all. Thus, CSIDNet manages to maintain the PSNR and SSIM values comparatively with faster run-time in comparison to others.

4.4 Qualitative comparison

Figures 2, 3, and 4 show the visualization of dehazed images from SOTS, HSTS, and natural hazy scenes, respectively. CSIDNet produces dehazed images without any visual artifacts. BCCR and MSCNN alter the color information near the sky region as they mainly focus on

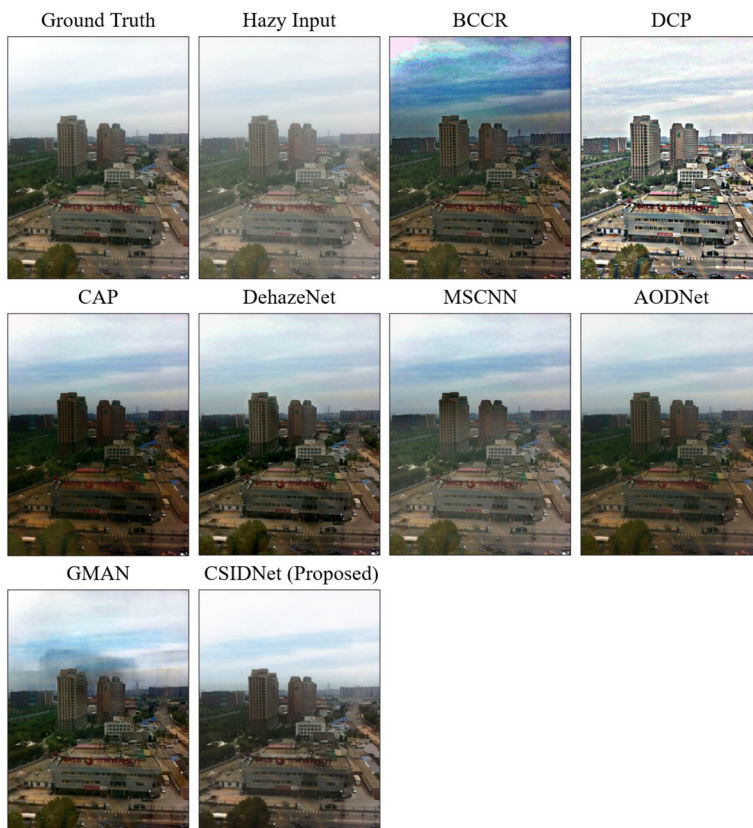


Fig. 3 Visual comparison on an image from HSTS of RESIDE dataset [17]



Fig. 4 Visual comparison on natural hazy image

Table 3 Performance measures for CSIDNet with different number of layers

Number of Layers	Testing Set			
	SOTS [17]		HSTS [17]	
	PSNR	SSIM	PSNR	SSIM
2	29.049	0.883	28.856	0.849
3	28.995	0.887	28.975	0.859
4	29.008	0.884	28.764	0.854
5	28.850	0.862	28.660	0.844
6	28.770	0.853	28.472	0.818
7	28.551	0.841	28.295	0.801
8	28.606	0.834	28.328	0.792

The bold entries are used to show the final selected number of layers in the proposed network

increasing the contrast. DCP generates dehazed images with halo artifacts near edges and fails to deal with the sky regions. CAP produces over saturated dehazed images and alters the color information. However, with respect to aforementioned methods, DehazeNet produces visually appealing dehazed images but is left with some haze. The dehazed images obtained using GMAN result in distortions which are easily visible in the visual comparison. In comparison among all state-of-the-art methods, the dehazed images obtained using AODNet seem better than others. However, the blur generated near edges distorts the textural information. CSIDNet gives visually pleasing results without any visible artifacts; on the other hand, nearly all the methods generate noticeable distortions, especially in the sky region. It is due to the possibility of excessive dehazing by other methods in regions with fine or light haze.

4.5 Discussions

The compact size of the proposed network has been experimented on different number of layers and number of filters/depth. Table 3 tabulates the values of PSNR and SSIM index obtained using trained models generated with different number of layers. It depicts that the performance on test datasets is better for 3 layers. Similarly, Table 4 tabulates the values of PSNR and SSIM index obtained using trained models generated with different number of filters. Figure 5 shows the plots with MSE as a loss function for different number of layers and different number of filters. Finally, 16 filters were chosen for the proposed network as they take lesser memory and calculation time without compromising the accuracy. Increasing the number of filters leads to increased memory requirements, which will be considerably less in the proposed network. Hence, a more compact structure is obtained.

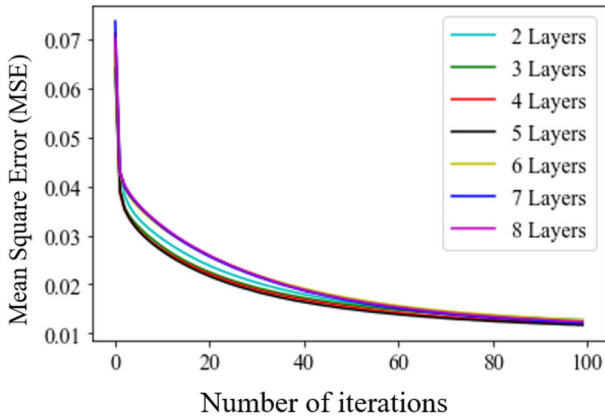
CSIDNet has been trained for five sets separately. Each set contains 200 images for training i.e. Set 1, Set 2, Set 3, Set 4, and Set 5. For quantitative comparison, the testing has been performed on images from SOTS and HSTS datasets in terms of PSNR and SSIM index measures. Table 5 shows the performance measures obtained with MSE loss function. The table tabulates the performance measures using trained models generated with training sets 1, 2, 3, 4, and 5. Similarly, Tables 6, 7, and 8, show the performance measures obtained with $L1$ loss function for 100, 150, and 200 epochs, respectively.

It can be observed that using MSE as a loss function for training, gives better PSNR and SSIM index measures. This can be interpreted by the mathematical behaviour of MSE which strongly penalizes the difference between ground truth and predicted value by squaring the error, as compared to $L1$ loss which only considers the absolute difference. Figs. 5 and 6 show the plots for MSE as a loss function and $L1$ loss function, respectively.

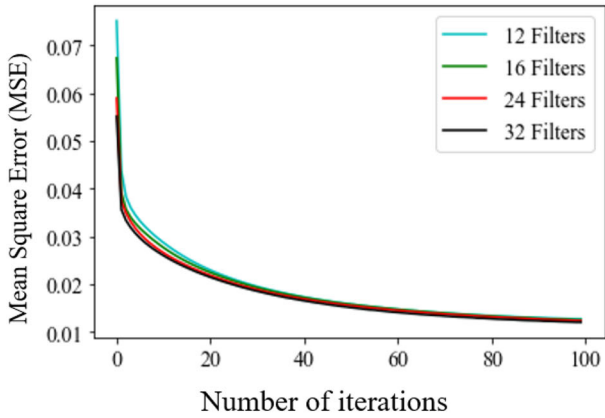
Table 4 Performance measures for CSIDNet with different number of filters

Number of Filters	Testing Set			
	SOTS [17]		HSTS [17]	
	PSNR	SSIM	PSNR	SSIM
12	28.999	0.885	28.765	0.853
16	29.045	0.887	28.975	0.856
24	29.056	0.888	28.912	0.857
32	29.033	0.887	28.892	0.855

The bold entries are used to show the final selected number of filters in the proposed network



(a)



(b)

Fig. 5 Plots for the mean square error with different (a) Number of layers and (b) Number of filters

Table 5 Performance measures for CSIDNet using MSE as loss function

Training Set	Testing Set			
	SOTS [17]		HSTS [17]	
	PSNR	SSIM	PSNR	SSIM
Set 1	29.045	0.887	28.975	0.856
Set 2	29.108	0.891	28.895	0.856
Set 3	28.982	0.880	28.838	0.844
Set 4	28.908	0.880	28.640	0.844
Set 5	28.932	0.878	28.791	0.842
Average	28.995	0.883	28.828	0.848

Table 6 Performance measures for CSIDNet using $L1$ loss function for 100 epochs

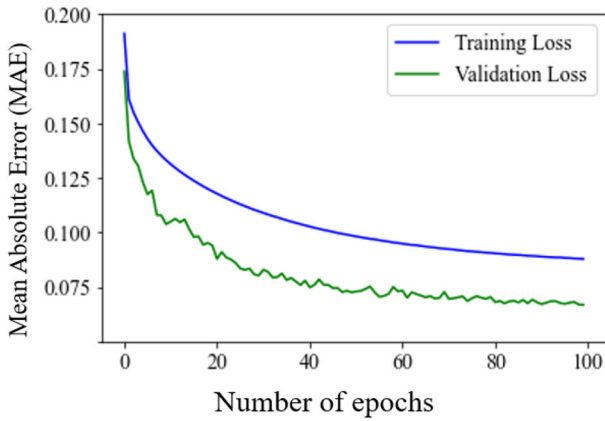
Training Set	Testing Set			
	SOTS [17]		HSTS [17]	
	PSNR	SSIM	PSNR	SSIM
Set 1	28.952	0.878	28.742	0.844
Set 2	28.960	0.878	28.727	0.842
Set 3	29.019	0.880	28.872	0.842
Set 4	28.791	0.878	28.600	0.843
Set 5	28.836	0.872	28.545	0.831
Average	28.912	0.877	28.697	0.840

Table 7 Performance measures for CSIDNet using $L1$ loss function for 150 epochs

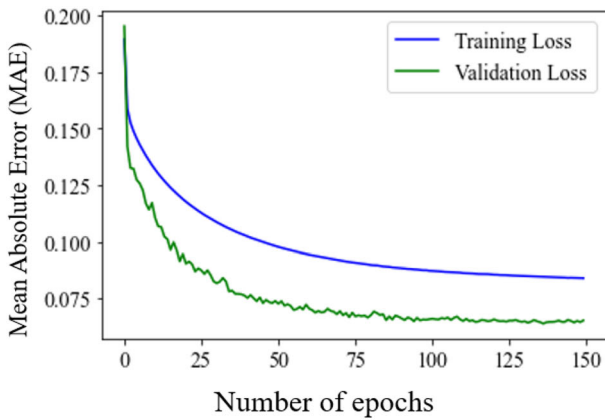
Training Set	Testing Set			
	SOTS [17]		HSTS [17]	
	PSNR	SSIM	PSNR	SSIM
Set 1	29.041	0.876	28.910	0.852
Set 2	29.067	0.885	28.937	0.852
Set 3	28.911	0.871	28.814	0.828
Set 4	28.805	0.878	28.577	0.840
Set 5	28.963	0.869	28.793	0.835
Average	28.957	0.876	28.806	0.841

Table 8 Performance measures for CSIDNet using $L1$ loss function for 200 epochs

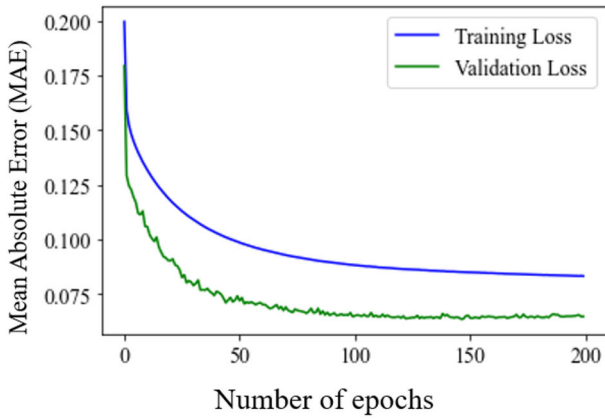
Training Set	Testing Set			
	SOTS [17]		HSTS [17]	
	PSNR	SSIM	PSNR	SSIM
Set 1	28.966	0.871	28.874	0.844
Set 2	29.062	0.885	28.889	0.855
Set 3	28.932	0.873	28.725	0.834
Set 4	28.959	0.880	28.751	0.847
Set 5	28.951	0.865	28.743	0.832
Average	28.974	0.875	28.796	0.842



(a)



(b)



(c)

Fig. 6 Plots for the mean absolute error/ $L1$ loss with (a) 100 epochs, (b) 150 epochs, and (c) 200 epochs

5 Conclusion

In this paper, a compact deep learning-based single image dehazing network named as *Compact Single Image Dehazing Network (CSIDNet)* has been proposed for outdoor scene enhancement. The proposed network not only outperforms several state-of-the-art dehazing models, but also sets a benchmark as a compact model with minimal resource requirements. The enhanced scene images obtained using CSIDNet successfully maintains trade off between speed and accuracy. The comparative analysis indicates that CSIDNet is faster to train with lesser run-time while maintaining its performance and robustness quantitatively and visually. The proposed network gives remarkable results, which suggests its scope in various critical and real-time applications. Moreover, the potential future scope is the development of an end-to-end network for image dehazing and denoising as well under non-uniform illumination conditions without considering any priors and assumptions.

References

1. Ancuti CO, Ancuti C (2013) Single image dehazing by multi-scale fusion. *IEEE Trans Image Process* 22(8):3271–3282. <https://doi.org/10.1109/TIP.2013.2262284>
2. Ancuti C, Ancuti CO (2014) Effective contrast-based dehazing for robust image matching. *IEEE Geosci Remote Sens Lett* 11(11):1871–1875. <https://doi.org/10.1109/LGRS.2014.2312314>
3. Burger W, Burge MJ (2016) *Digital image processing: an algorithmic introduction using java*. Springer, London
4. Cai B, Xu X, Jia K, Qing C, Tao D (2016) DehazeNet: An end-to-end system for single image haze removal. *IEEE Trans Image Process* 25(11):5187–5198. <https://doi.org/10.1109/TIP.2016.2598681>
5. Chaudhry AM, Riaz MM, Ghafoor A (2018) A framework for outdoor RGB image enhancement and dehazing. *IEEE Geosci Remote Sens Lett* 15(6):932–936. <https://doi.org/10.1109/LGRS.2018.2814016>
6. Fattal R (2008) Single image dehazing. *ACM Transactions on Graphics (TOG)* 27(3):1–9. <https://doi.org/10.1145/1360612.1360671>
7. Gonzalez RC, Woods RE (2006) *Digital image processing*, 3rd. Prentice-Hall, Inc. Upper Saddle River, New Jersey
8. He K, Sun J, Tang X (2010) Fast matting using large kernel matting Laplacian matrices. In: 2010 IEEE computer society conference on computer vision and pattern recognition, San Francisco, CA, USA, pp 2165–2172. <https://doi.org/10.1109/CVPR.2010.5539896>
9. He K, Sun J, Tang X (2011) Single image haze removal using dark channel prior. *IEEE Trans Pattern Anal Mach Intell* 33(12):2341–2353. <https://doi.org/10.1109/TPAMI.2010.168>
10. He K, Sun J, Tang X (2013) Guided image filtering. *IEEE Trans Pattern Anal Mach Intell* 35(6):1397–1409. <https://doi.org/10.1109/TPAMI.2012.213>
11. He K, Zhang X, Ren S, Sun J (2015) Delving deep into rectifiers: Surpassing human-level performance on imageNet classification. In: 2015 IEEE international conference on computer vision (ICCV), Santiago, Chile. <https://doi.org/10.1109/ICCV.2015.123>
12. He K, Zhang X, Ren S, Sun J (2016a) Identity mappings in deep residual networks. In: European conference on computer vision (ECCV) 2016, pp 630–645. https://doi.org/10.1007/978-3-319-46493-0_38
13. He K, Zhang X, Ren S, Sun J (2016b) Deep residual learning for image recognition. In: 2016 IEEE Conference on computer vision and pattern recognition (CVPR), Las Vegas, NV, USA. <https://doi.org/10.1109/CVPR.2016.90>
14. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: *NIPS'12 Proceedings of the 25th international conference on neural information processing systems - Volume 1*, Lake Tahoe, Nevada, vol 1, pp 1097–1105
15. Levin A, Lischinski D, Weiss Y (2008) A closed-form solution to natural image matting. *IEEE Trans Pattern Anal Mach Intell* 30(2):228–242. <https://doi.org/10.1109/TPAMI.2007.1177>
16. Li B, Peng X, Wang Z, Xu J, Feng D (2017) AOD-Net: All-in-one dehazing network. In: 2017 IEEE International conference on computer vision (ICCV), Venice, Italy, pp 4780–4788. <https://doi.org/10.1109/ICCV.2017.511>
17. Li B, Ren W, Fu D, Tao D, Feng D, Zeng W, Wang Z (2019) Benchmarking single image dehazing and beyond. *IEEE Trans Image Process* 28(1):492–505. <https://doi.org/10.1109/TIP.2018.2867951>

18. Liu Z, Xiao B, Alrabeiah M, Wang K, Chen J (2019) Single image dehazing with a generic model-agnostic convolutional neural network. *IEEE Signal Process Lett* 26(6):833–837. <https://doi.org/10.1109/LSP.2019.2910403>
19. Long J, Shi Z, Tang W, Zhang C (2014) Single remote sensing image dehazing. *IEEE Geosci Remote Sens Letters* 11(1):59–63. <https://doi.org/10.1109/LGRS.2013.2245857>
20. Lu X, Ma C, Ni B, Yang X, Reid I, Yang MH (2018) Deep regression tracking with shrinkage loss. In: *Proceedings of the European conference on computer vision (ECCV)*, pp 353–369
21. Lu X, Ni B, Ma C, Yang X (2019) Learning transform-aware attentive network for object tracking. *Neurocomputing* 349:133–144. <https://doi.org/10.1016/j.neucom.2019.02.021>
22. Lu X, Wang W, Ma C, Shen J, Shao L, Porikli F (2019) See more, know more: unsupervised video object segmentation with co-attention siamese networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 3623–3632
23. Meng G, Wang Y, Duan J, Xiang S, Pan C (2013) Efficient image dehazing with boundary constraint and contextual regularization. In: *IEEE international conference on computer vision (ICCV)*, Sydney, NSW, Australia, pp 617–624
24. Narasimhan SG, Nayar SK (2000) Chromatic framework for vision in bad weather. In: *Proceedings IEEE conference on computer vision and pattern recognition, CVPR 2000 (Cat. No.PR00662)*, Hilton Head Island, SC, USA, vol 1, pp 598–605. <https://doi.org/10.1109/CVPR.2000.855874>
25. Narasimhan SG, Nayar SK (2002) Vision and the atmosphere. *Int J Comput Vis* 48(3):233–254. <https://doi.org/10.1023/A:1016328200723>
26. Narasimhan SG, Nayar SK (2003) Contrast restoration of weather degraded images. *IEEE Trans Pattern Anal Mach Intell* 25(6):713–724. <https://doi.org/10.1109/TPAMI.2003.1201821>
27. Qin X, Wang Z, Bai Y, Xie X, Jia H (2020) FFA-Net: Feature fusion attention network for single image dehazing. In: *AAAI*, pp 11908–11915
28. Reinhard E, Adhikhmin M, Gooch B, Shirley P (2001) Color transfer between images. *IEEE Comput Graph Appl* 21(5):34–41. <https://doi.org/10.1109/38.946629>
29. Ren W, Liu S, Zhang H, Pan J, Cao X, Yang MH (2016) Single image dehazing via multi-scale convolutional neural networks. *European Conference Comput Vis* 9906:154–169. https://doi.org/10.1007/978-3-319-46475-6_10
30. Ren W, Ma L, Zhang J, Pan J, Cao X, Liu W, Yang MH (2018) Gated fusion network for single image dehazing. In: *2018 IEEE/CVF conference on computer vision and pattern recognition, Salt Lake City, UT, USA*, pp 1–9. <https://doi.org/10.1109/CVPR.2018.00343>
31. Rosolia U, Bruyne SD, Alleyne AG (2017) Autonomous vehicle control: a nonconvex approach for obstacle avoidance. *IEEE Trans Control Syst Technol* 25(2):469–484. <https://doi.org/10.1109/TCST.2016.2569468>
32. Schechner YY, Narasimhan SG, Nayar SK (2001) Instant dehazing of images using polarization. In: *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, Kauai, HI, USA. <https://doi.org/10.1109/CVPR.2001.990493>
33. Tan RT (2008) Visibility in bad weather from a single image. In: *2008 IEEE conference on computer vision and pattern recognition, Anchorage, AK, USA*, pp 1–8. <https://doi.org/10.1109/CVPR.2008.4587643>
34. Tang K, Yang J, Wang J (2014) Investigating haze-relevant features in a learning framework for image dehazing. In: *2014 IEEE Conference on computer vision and pattern recognition, Columbus, OH, USA*, pp 2995–3002. <https://doi.org/10.1109/CVPR.2014.383>
35. Tarel JP, Hautière N (2009) Fast visibility restoration from a single color or gray level image. In: *2009 IEEE 12th international conference on computer vision, Kyoto, Japan*, pp 2201–2208. <https://doi.org/10.1109/ICCV.2009.5459251>
36. Wang A, Wang W, Liu J, Gu N (2019) AIPNet: Image-to-image single image dehazing with atmospheric illumination prior. *IEEE Trans Image Process* 28(1):381–393. <https://doi.org/10.1109/TIP.2018.2868567>
37. Yang X, Li H, Fan YL, Chen R (2019) Single image haze removal via region detection network. *IEEE Trans Multimed* 21(10):2545–2560. <https://doi.org/10.1109/TMM.2019.2908375>
38. Zhu Q, Mai J, Shao L (2015) A fast single image haze removal algorithm using color attenuation prior. *IEEE Trans Image Process* 24(11):3522–3533. <https://doi.org/10.1109/TIP.2015.2446191>