Dehazing Network based on Haze Density

Yanling Hua[1], Zhengrong Zuo[1*]
[1]National Key Laboratory of Science and Technology on Multispectral Information Processing, Huazhong University of Science and Technology, Wuhan 430074,China

ABSTRACT

Single image dehazing is a challenging ill-posed restoration problem. Most of dehazing algorithms follow the classical atmospheric scattering model and adopt same parameters for different hazy density areas in hazy images. In this paper, we proposed an end-to-end dehazing algorithm, called Dehazing Network based on Haze Density(DNBHD). The proposed network involves a haze density map estimation network and a dehazing network. By the estimated haze density map, hazy image is divided into a mist region and a dense fog region which are respectively feed into dehazing network. Compared with previous dehazing algorithm, DNBHD is independent on the atmospheric scattering model, and considers uniform fog distribution in images. We use different parameters to handle different hazy density regions, avoiding color distortion and inappropriate brightness caused by overall defogging. The experiments show our algorithm achieves significant improvements over the state-of-the-art methods.
Keywords: Single image dehazing, Haze density, Transmission map estimation, End-to-end defogging

1. INTRODUCTION

Single image dehazing is a challenging ill-posed image restoration problem. The images collected under hazy conditions are prone to characteristic problems such as color distortion and offset, small dynamic compression range, and low contrast. Moreover, hazy images not only fail to achieve the desired visual effect, but also increase the difficulty of object detection and recognition.

In order to effectively process hazy images, an atmospheric scattering model[1] is proposed according to the imaging mechanism of hazy images, as shown in Equations 1 and Equations 2:

$$I(x) = J(x)t(x) + A(x)(1 - t(x)) \tag{1}$$

$$t(x) = e^{-\beta d(x)} \tag{2}$$

where I(x) is hazy image, J(x) is original clean image; t(x) is the transmission map; A(x) is the atmospheric light. Single image dehazing algorithms are mainly divided into two main groups, prior-based[4][5][6][5][6][7] and learning-based[12][9][10][11][12][13][14] methods. The prior-based defogging algorithms need to manually design image features. Their performance depends on the accuracy of the prior knowledge. So they are difficult to have a good generalization for all types of scenes. Most of learning-based networks rely on the atmospheric scattering model which needs to estimate transmission map and atmospheric light value respectively. After that, they restore the clear map according to Equation 1. These methods have two limitations. First, the inaccurate estimation of the transmission map and atmospheric light will affect the overall dehazing performance and image sharpness of the defogged image. Second, the linear model is not enough to simulate the complex fogging mechanism. Another type of learning-based defogging algorithms directly recovers the clear map from the foggy image, using network to learn complex transformation functions. However, due to the lack of physical model and the nature of parameter sharing of convolutional neural network, the same parameters are used in different hazy density regions in hazy image. It leads to partial darkness or excessive brightness after defogging, and leaving some fog in images.

The transmission map represents the proportion of the reflected light of scene that can reach the detection system through the particle attenuation, and it is negatively correlated with the hazy density. So we can obtain haze density map by estimating transmission map. The proposed Dehazing Network based on Haze Density (DNBHD) is an end-to-end dehazing algorithm based on haze density. Figure 2 shows the overview of DNBHD. First, haze density map estimation network estimates the transmission map of the input hazy image. The transmission map is inversely proportional to haze density, so the region where the luminance is small in the transmission map has thick fog. According to the transmission

***Corresponding author:** E-mail: zhrzuo@hust.edu.cn

map, hazy image is divided into a dense fog map and a mist map. Then, the obtained two fog maps are connected in channels and feed into dehazing network. Finally, the first three channels and the last three channels of the output of dehazing network are combined at the pixel level and feed into output branch. The experiments show that the proposed algorithm avoids the color distortion and inappropriate brightness caused by the overall defogging, and gets more detailed results.

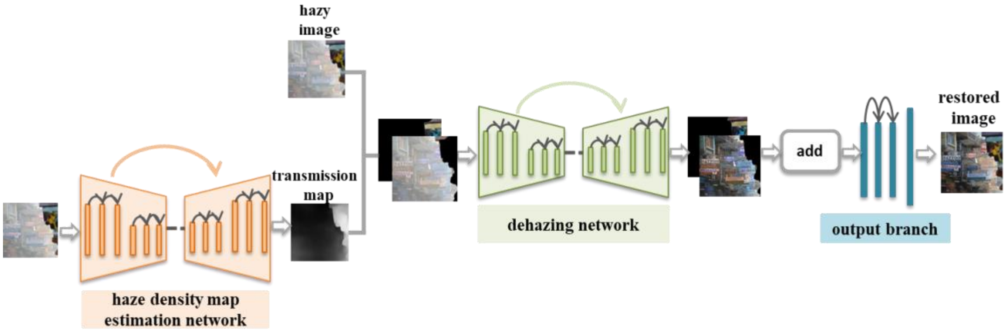


Figure 1: Overview of proposed Dehazing Network based on Haze Density (DNBHD)

The contributions of this paper are two-fold:

(1) A new defog network based on haze density is proposed. Unlike previous defog networks based on linear atmospheric scattering models, DNBHD learns complex transformation functions from hazy image to clear image through the network.

(2) In order to make the network have different perceptions of different haze density regions in the image, we divide the hazy image into two regions, dense fog region and light fog region, by the estimated transmission map, and respectively restore the clear image.

2. RELATED WORK

In this section, we will review the algorithms for single image dehazing, including prior-based image defogging algorithms and learning-based image defogging algorithms.

Prior-based image dehazing algorithm. Prior-based image dehazing algorithm recovers clean map according to the atmospheric scattering model, mainly aim at estimating transmission map. He Kaiming et al[4] proposed the dark channel prior(DCP) algorithm. They found that the dark channel value of the local area of the clean map is almost close to zero. According to this prior, the transmission map is obtained by calculating the dark channel map of the fog map. Zhu et al.[5] proposed the color attenuation prior(CAP) algorithm to recover depth information. In this method, they used a linear model to obtain depth map by observing the depth is proportional to the difference between saturation and brightness. Berman.et al[6] proposed a non-local patch prior that colors of a haze-free image could be well approximated by a few hundred distinct color clusters in RGB space. The authors assumed that due to the presence of fog, the color clusters are in the form of lines. They restored depth information based on this prior.

Learning-based defogging algorithm. The learning-based dehazing algorithm mainly uses convolutional neural networks to recover clear images. Ren et al.[9] proposed Multi-Scale CNN(MSCNN). The MSCNN contains coarse-scale and fine-scale networks to obtain a transmission map. Fine-scale networks refine the transmission map obtained from the coarse-scale network. Li et al.[10] proposed an end-to-end dehazing network, the All-In-One Dehazing Network (AOD-net). Through the reformed atmospheric scattering model, the atmospheric light value and transmission map estimation are put into a network to recover the clear map. Kangfu Mei al.[11] proposed a defogging network based on encoder and decoder network, Progressive Feature Fusion Network(PFF-net). Recently, Bharath Raj N.al. et al.[12] proposed a dehazing algorithm based on conditional generation against networks. The algorithm uses the encoder decoder structure as the generator and the patch gan as the discriminator.

3. PROPOSED METHOD

In the hazy image, the haze density differs in different regions. Generally, the fog density in the near-field of the image is lower than that in the distant region. As for the whole image defogging, it may lead to over enhancement in the near-field and incomplete defogging in the far-field. In view of this, we propose a dehazing network based on haze density. According to the estimated transmission map, the foggy image is divided into different hazy density regions, and each region is separately defogged.

3.1 Haze density map estimation network

In order to obtain a more accurate transmission map, we use the U-net structure. Compared with encoder-decoder structure, U-net directly connects the low-level feature map with the high-level feature map. In this way, the detailed information of feature maps in different resolutions is better preserved. At the same time, in order to achieve the maximum flow of network information, we use dense-block as the basic module of the network. Each layer in the dense-block is connected to each other. The upper part of the U-net can be thought of as an encoder that maps the high-level representation to the low-level representation by down-sampling. And it provides the corresponding resolution maps for the decoder. The decoder reconstructs the encoded low-level representation to the high-level representation by up-sampled to get the transmission map. The structure of the network is shown in Figure 2. It consists of densely connected convolutional blocks (CBs), down-sampled convolutional blocks (CDs), and up-sampled convolutional blocks (CUs). The CB module consists of three submodules, each of which comprises of {bn-Relu-Convolution} operations, and each submodule is connected end to end. The CD module is an up-sampled module and consists of {bn-conv-dropout-pool}. The CU module is a transpose convolution operation. The kennel size of all convolutional layers is 3.

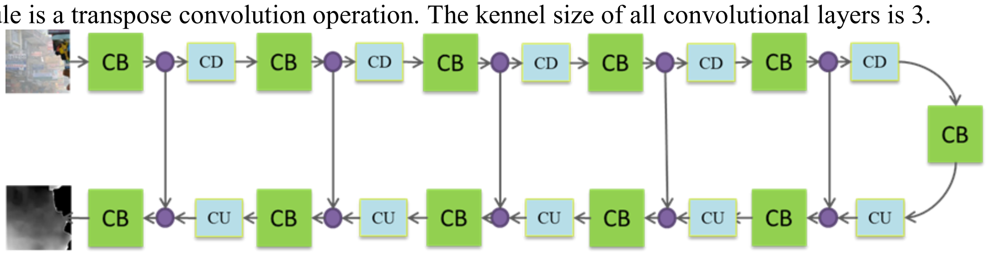


Figure 2: Transmission map estimation network.

3.2 Dehazing network

The dehazing network consists of three modules: hazy image segmentation module, hazy image defogging module, and a pixel level merge module.

Hazy image segmentation module. After getting the transmission map corresponding to the hazy image, we spatially segment the transmission map. The transmission map is inversely related to the fog concentration, so we set the grayscale threshold w based on the value of the transmission map and divide the transmission map into two parts. According to the obtained two transmission maps, the hazy image is divided into two parts: a dense fog map (a haze map with a transmission map smaller than w) and a mist map (a haze map with a transmission map larger than w). For the determination of threshold w, we conducted experiments on 5 w with 0.5 as the center within the range of 0-1 with 0.1 as the interval. Then we tested 5 trained networks on the open data set and selected the best w=0.6. Figure 3 shows the results of the transmission and the segmentation of the fog map.

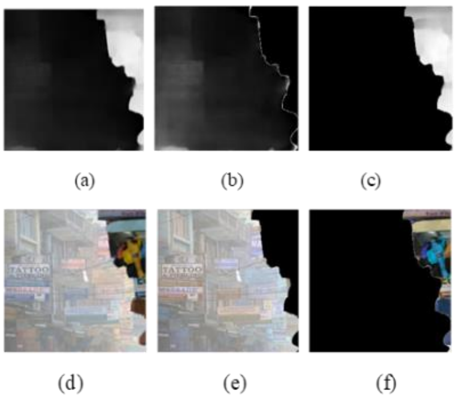


Figure 3. The results of the transmission and the segmentation of the fog map. (a): Transmission map. (b): Transmission map while is lower than w=0.6. (c): Transmission map while is greater than w=0.6. (d):Hazy image. (e):Dense fog map. (f) mist map

Hazy image defogging module. After getting dense fog map and mist map, we connect them in the channel level. So, the input of the dehazing network has six channels. In this way, the network have different perceptions of different haze density regions in the image. The structure of dehazing network has the same structure as the transmission estimation network. And the output convolution layer is changed to 6 corresponding to the input of dehazing net.

Pixel-level merge module. The first three channels and the last three channels of the output of dehazing network are added at the pixel level and get three channels image.

3.3 Output branch

The output branch is used to further optimize the result. The output branch consists of a densely connected block and a convolutional layer. The densely connected block module consists of three submodules, each of which comprises of {bn-Relu-Convolution} operations, and each submodule is connected end to end. The kennel size of all convolutional layers is 3. The output channels of the convolutional layer is 3 and its kennel size is 1.

3.4 Loss function

The proposed defog network mainly accomplishes two tasks, one is the estimation of the transmission map, and the other is the defogging. So the loss function is divided into two parts: the loss function of the estimated transmission map and the defogging loss function.

Loss functions of estimating transmission map. Since L2 loss tends to blur the final result, we use L1 loss in this paper, defined as follows

$$L_{l1\_t} = \sum_{w,h,c} ||T(I) - t||_1 \tag{3}$$

where T(*) is the transmission map estimation network, I is the hazy image, t is the target transmittance map, and w, h, c are the dimensions of the output feature map. At the same time, in order to optimize the edge information of the transmission map, we introduce the perceptual loss function [15][16] L_{vgg_t}, defined as follows

$$L_{vgg\_t} = \sum_{w,h,c} ||V(T(I)) - V(t)||_2 \tag{4}$$

where V(*) is the output of the Pool-4 layer of the VGG-19 network, and w, h, and c are the length, width, and channels of the feature map.

Image defogging loss functions. The loss function of the defogging consists of the L1 loss function L_{l1_I} and the perceptual loss function L_{vgg_I}, defined as follows

$$L_{l1\_t} = \sum_{w,h,c} ||S - C||_1 \tag{6}$$

$$L_{vgg\_t} = \sum_{w,h,c} ||V(S) - V(C)||_2 \tag{7}$$

where $D(*)$ denotes dehazing network, S is the dehazed image, and C is the target image.

4. EXPERIMENTAL RESULTS

4.1 Datasets

Training datasets. Similar to the existing deep learning-based defog methods, we synthesize the training samples {hazy/clean/transmission map} based on (1). Most of these methods use NYU_depth datasets[17], which are all indoor images, but the main application scene of the defogging algorithm is outdoor. So we propose a new synthetic datasets based on ReDWeb_V1[18]. The ReDWeb V1 dataset consists of 3.6K RGB-RD images, covering both indoor and outdoor scenes. Given the clear image J and the depth map d, we synthesis 10k hazy image and the transmission map according to the atmospheric scattering model. We get the atmospheric light values random sampled from (0.8, 1) and atmospheric scattering coefficient β sampled from (0.4, 1.6). Finally, the size of the clear map and the synthesized fog map are cropped to 256×256.

Test database. We use the SOTS dataset in the public dataset RESIDE[19] as the algorithm performance test dataset, which consists of 500 outdoor fog maps and clear maps, not overlapping with the training database.

4.2 Training details

We adopt Adamoptimizer with beta1=0.5, beta2=0.999 and rate=10-8. The learning rate is 0.001. All trainable variables are initialized using Xavier method. We take ten percent of training data as verification dataset. In order to speed up the training, we first train transmission estimation network, followed by training the dehazing network, and finally fine-tuning the whole network.

4.3 Ablation study

Determination of threshold w. We use peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) to evaluate the dehazing performance of the algorithm. In order to compare the effects of different thresholds w, we evaluate the results of the networks trained at different thresholds on the public test dataset under the same settings. We first trained a defog network with a threshold of 0.5. On the basis of this net, we modified w and fine-tuned the network parameters. In this way, the training time is greatly reduced. We finally took w=0.6. The results are shown in Table 1.

Table 1: Algorithm performance under different w: Score=PSNR/20+SSIM

w	PSNR	SSIM	Score
0.3	22.838	0.913	2.0549
0.4	23.674	0.915	2.0987
0.5	24.664	0.929	2.1622
0.6	24.748	0.935	2.1729
0.7	23.688	0.923	2.1074

Loss function of estimate transmission map. To illustrate the effectiveness of the loss function of estimating transmission map, we experimented with different loss functions for the same network structure: (i) L1 loss, (ii) L1 loss + perceptual loss. The experiment results are shown in Figure 4. We can see that the transmission map obtained by using the l1 loss function and the perceptual loss function is much more smooth and better refine the edges.

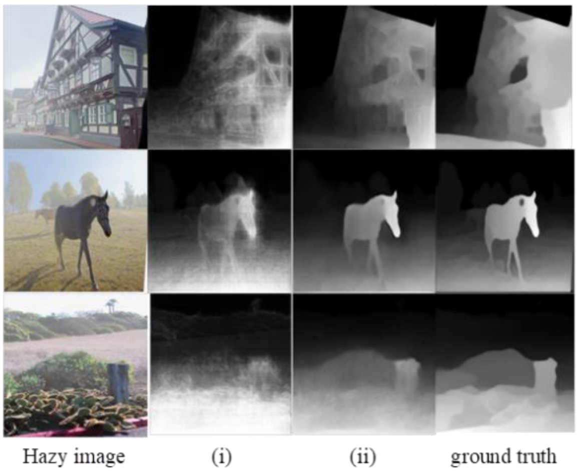


Figure 3. Estimated transmission map under different loss function.

Structure and loss function of dehazing. We conduct three experiment about dehazing network: (a) dehazing network without Haze density map estimation network, (aa) dehazing network without perceptual loss, (aaa) proposed network. Table2 shows the average PSNR and SSIM results on SOTS.

Table 2: Average PSNR and SSIM results for different configurations of the proposed network

w	PSNR	SSIM	Score
(a)	23.781	0.895	2.0837
(aa)	22.813	0.903	2.0437
(aaa)	24.748	0.935	2.1729

4.4 Comparison with state-of-the-art methods

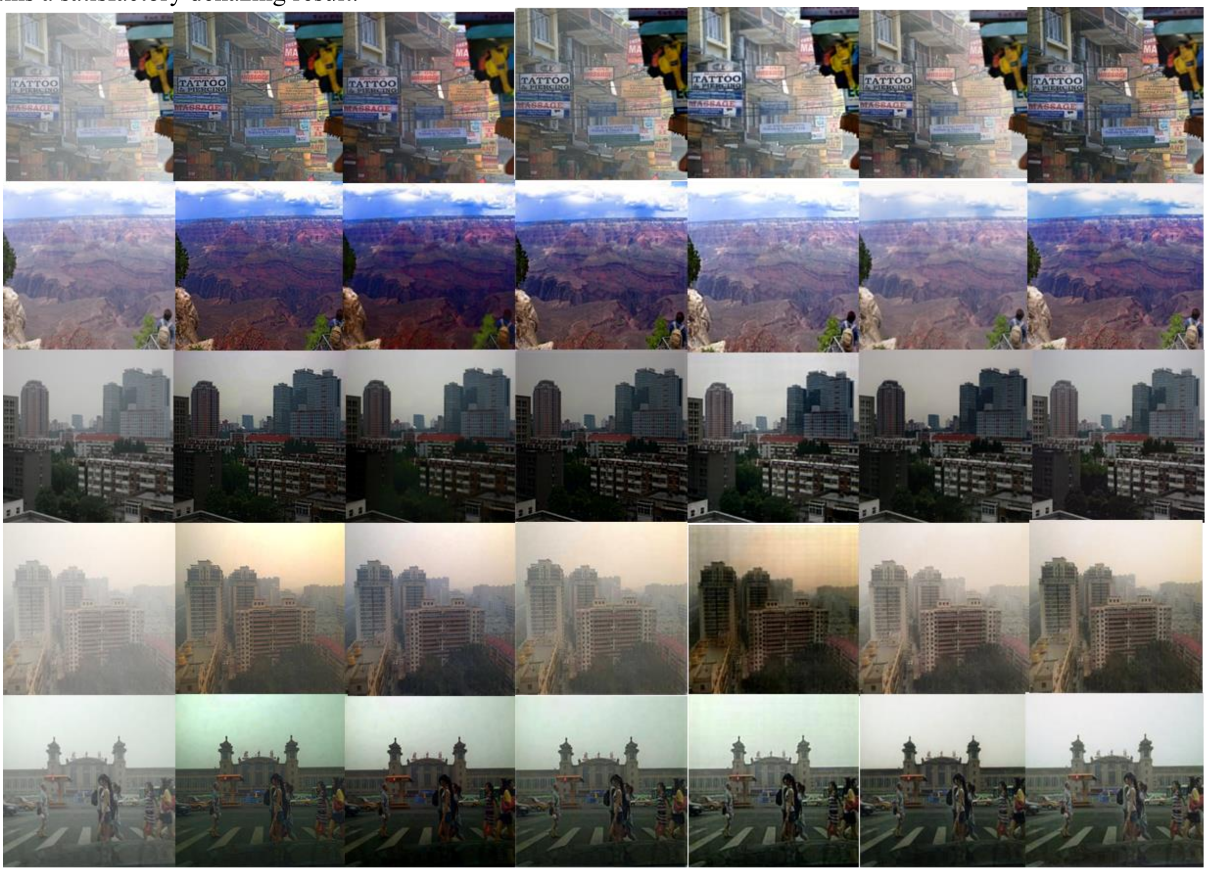
Quantitative Results. To verify the generalization capabilities of the model, we tested the model performance on the public dataset SOTS. The DNBHD is compared with other advanced dehazing algorithms. The comparison results are shown in Table 1. The results show that the performance of our algorithm exceeds other algorithms.

Table 2: Results of the quantitative analysis conducted on SOTS

Model	PSNR	SSIM
DCP[0]	19.155	0.899
CAP[5]	18.481	0.773
AOD-net[10]	20.292	0.875
PFF-net[11]	21.252	0.839
Dehaze-GAN[12]	23.595	0.909
DNBHD (proposed)	24.748	0.935

Qualitative Results. Figure 3 shows five real scenes hazy images and the corresponding dehazing results generated by several algorithms. It can be seen from the first and second lines that the result of the AOD-net[10], PFF-net[11] and

Dehaze-GAN[12] retain more fog, and the DCP[4] and CAP[5] makes the local area too dark. In the third row, the defogging results of DCP[4], AOD[10], Dehaze-GAN[12] lost information in the building and the forest, in stark contrast to the results of DNBHD. In the fourth line and last line, DNBHD better restores the color information of the sky and obtains a satisfactory dehazing result.



Input Image DCP CAP AOD PFF Dehaze-GAN Proposed

Figure 3: Dehazing results evaluated on real-world images

5 CONCLUSION

In this paper, a new end-to-end de-fog network is proposed. By estimating the transmission map, hazy image is divided into a mist region and a dense fog region, which are respectively placed in dehazing network. Finally, pixel-level merging is adopted to obtain a complete and clear image. Compared with the previous dehazing algorithm, DNBHD does not depend on the atmospheric scattering model, and considers the uniform distribution of fog in the image. Different networks are used to handle different hazy density regions, avoiding the color distortion caused by the overall defogging and inappropriate brightness while restoring more details. The experiments show that our algorithm achieves significant improvements over the state-of-the-art methods and can apply to more scenes.

ACKNOWLEDGEMENT

This work was supported by National Natural Science Foundation of China (Grant No. 61773389).

REFERENCE

[1] Cantor, A.: Optics of the atmosphere–scattering by molecules and particles. IEEE Journal of Quantum Electronics 14 (1978) 698–699
[2] J.-P. Tarel and N. Hautiere. Fast visibility restoration from a single color or gray level image. In IEEE International Conference on Computer Vision, pages 2201–2208. IEEE, 2009.
[3] C. O. Ancuti, C. Ancuti, C. Hermans, and P. Bekaert. A fast semi-inverse approach to detect and remove the haze from a single image. In Asian Conference on Computer Vision, pages 501–514. Springer, 2010.
[4] He, Kaiming, Jian Sun, and Xiaoou Tang. "Single image haze removal using dark channel prior." IEEE transactions on pattern analysis and machine intelligence 33.12 (2011): 2341-2353.
[5] Q. Zhu, J. Mai, and L. Shao. A fast single image haze removal algorithm using color attenuation prior. IEEE Transactions on Image Processing, 24(11):3522–3533, 2015.
[6] D. Berman, S. Avidan, et al. Non-local image dehazing. In CVPR, pages 1674–1682, 2016. 2, 3, 6, 7, 8
[7] C. Ancuti, C. O. Ancuti, C. De Vleeschouwer, and A. C. Bovik.Night-time dehazing by fusion. In IEEE International Conference on Image Processing (ICIP), pages 2256–2260. IEEE, 2016.
[8] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao. Dehazenet: An end-to-end system for single image haze removal. IEEE Transactions on Image Processing,25(11):5187–5198, 2016.
[9] Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., Yang, M.H.: Single image dehazing via multi-scale convolutional neural networks. In: Proceedings of European Conference on Computer Vision, Springer (2016)
[10] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng. An all-in-one network for dehazing and beyond. ICCV, 2017.
[11] Kangfu Mei, Aiwen Jiang, Juncheng Li2[, and Mingwen Wang. Progressive Feature Fusion Network for Realistic Image Dehazing. arXiv:1810.02283, 2018.
[12] Bharath Raj N., Venkateswaran N. Single Image Haze Removal using a Generative Adversarial Network. arXiv:1810.09479, 2018
[13] K. Swami and S. K. Das. Candy: Conditional adversarial networks based fully end-to-end system for single image haze removal. arXiv preprint arXiv:1801.02892, 2018.
[14] H. Zhang, V. Sindagi, and V. M. Patel. Joint transmission map estimation and dehazing using deep networks. arXiv preprint arXiv:1708.00581, 2017.
[15] Ledig, Christian, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken et al. "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4681-4690. 2017.
[16] Zhang He, Vishwanath Sindagi, and Vishal M. Patel. "Image de-raining using a conditional generative adversarial network." arXiv preprint arXiv:1701.05957 (2017).
[17] Silberman, Nathan, Derek Hoiem, Pushmeet Kohli, and Rob Fergus. "Indoor segmentation and support inference from rgbd images." In European Conference on Computer Vision, pp. 746760. Springer, Berlin, Heidelberg, 2012.
[18] Ke Xian, Chunhua Shen, Zhiguo Cao, Hao Lu, Yang Xiao, Ruibo Li, Zhenbo Luo. Monocular Relative Depth Perception with Web Stereo Data Supervision. IEEE Conf. Computer Vision and Pattern Recognition (CVPR) 2018
[19] Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W., Wang, Z.: Benchmarking single image dehazing and beyond. IEEE Transactions on Image Processing (2018)