Huazhong University of Science and Technology

Dehazing Algorithm Based on Image Fogging Degree

Authors:
Yanling Hua

Instructors:
Zhengrong Zuo

January 9, 2025

Contents

1 Introduction

To effectively process foggy images, McCartney proposed the atmospheric scattering model [1], described as follows:

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

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

where $I(x)$ represents the hazy image; $J(x)$ is the original, non-degraded image to be recovered; $t(x)$ denotes the transmission map corresponding to the hazy image; A is the atmospheric light value; $d(x)$ is the scene depth; and β is the haze density parameter.

Single-image dehazing algorithms are categorized into two types: dehazing algorithms based on prior knowledge and dehazing algorithms based on deep learning. Prior-based dehazing algorithms require manual design of image features, and their performance depends on the accuracy of the preceding knowledge, leading to poor generalization capability.

Learning-based dehazing algorithms can be further divided into two categories depending on whether the atmospheric scattering model is used to design the network structure. For algorithms that rely on the atmospheric scattering model, the transmission map and atmospheric light value are estimated separately, and the clear image is restored using Equation (2-1). This approach has two limitations: first, inaccuracies in estimating the transmission map and atmospheric light value affect the degree of dehazing; second, the linear model is insufficient to simulate the complex haze formation mechanism.

The other type of learning-based dehazing algorithm directly restores a clear image from the hazy image by learning the complex transformation function from hazy to clear images through the network. However, due to the lack of constraints from physical models and the parameter-sharing nature of convolutional neural networks, regions with varying haze densities in the hazy image are processed using the same parameters. This leads to local darkening or over-brightening of the dehazed image, resulting in incomplete haze removal.

The transmission map of a hazy image reflects the proportion of scene-reflected light that can reach the detection system after attenuation by particles, which is negatively correlated with the degree of haze in the image. Based on the transmission map, this paper designs a dehazing convolutional network capable of processing regions with varying haze densities in parallel.

2 Related Work

Haze-induced image degradation has adversely impacted both daily life and advanced computer vision tasks. Consequently, the study of image dehazing algorithms holds significant practical value and application potential. These algorithms can be broadly categorized into two types based on their principles: prior knowledge-based algorithms and deep learning-based algorithms.

2.1 Dehazing Algorithms Based on Prior Knowledge

Dehazing algorithms based on prior knowledge restore images according to the atmospheric scattering model, with the primary task being the estimation of the transmission map. He et al. proposed the Dark Channel Prior (DCP) algorithm, where the authors observed that the dark channel values in clear image patches are almost zero [2]. Based on this prior, the transmission map is obtained by calculating the dark channel of the hazy image. Zhu et al. introduced the Color Attenuation Prior (CAP) algorithm to recover depth information [3]. In this method, the transmission map is computed by observing the proportional relationship between depth information and the difference in image saturation and brightness. Berman et al. proposed a non-local image dehazing method, assuming that the colors of haze-free images can be approximated by a few hundred distinct color clusters in the RGB space [4]. They hypothesized that the presence of haze transforms these clusters into line structures and used this prior to recover depth information and compute the transmission map.

2.2 Dehazing Algorithms Based on Deep Learning

Deep learning-based dehazing algorithms mainly utilize convolutional neural networks to restore clear images. Ren et al. proposed the Multi-Scale Dehazing Network (MSCNN), which includes coarse and fine-scale networks to estimate the transmission map [5]. The fine-scale network refines the transmission map obtained from the coarse-scale network. Li et al. introduced an end-to-end dehazing network (AOD-Net) that recovers clear images within a single network by transforming the calculation of the atmospheric scattering model [6]. Kangfu Mei et al. proposed a dehazing network based on an encoder-decoder structure, named Progressive Feature Fusion Network (PFF-Net), which directly learns the mapping from hazy to haze-free images [7]. Bharath Raj N. et al. developed a dehazing algorithm using a conditional generative network (Dehaze-GAN). This algorithm employs an encoder-decoder structure as the generator to learn the mapping from hazy to haze-free images and utilizes PatchGAN as the discriminator [8]. Malav et al. introduced an end-to-end dehazing and smoke-removal network (DHSGAN), which is trained under a generative adversarial network framework. The transmission map and the hazy image are combined along channels as the input to the generator to restore haze-free images [9].

From the above methods, we can observe that significant progress has been made in image dehazing algorithms within both traditional image processing and deep learning fields. However, existing algorithms exhibit weak generalization when dealing with hazy images with uneven haze distribution or real-world images from diverse scenes. Therefore, designing more targeted network structures to address the phenomenon of uneven haze distribution in images is key to improving the performance of dehazing algorithms.

3 Method

3.1 Dehazing Network

The dehazing model based on haze degree prediction proposed in this paper is illustrated in Figure 1. The model comprises four modules: a feature encoding module, a transmission map estimation module, a dehazing module, and a separation dehazing module. The feature encoding module extracts multi-scale image features, which are then passed to the corresponding convolutional blocks of the transmission map estimation module and the dehazing module. The transmission map and image features extracted by the dehazing module are input into the separation dehazing module to restore the clear image.

Different regions of hazy images exhibit varying degrees of haze. Typically, the haze degree in near-field regions is lower than in far-field regions. Global dehazing might lead to over-enhancement in near-field regions and incomplete dehazing in far-field regions. Based on this observation, a separation dehazing module is constructed. The basic idea of the separation convolution module is to utilize the transmission map estimated by the transmission map estimation module to determine the haze degree. By setting an appropriate threshold, weak-haze and strong-haze masks can be calculated, guiding the convolution process of the separation convolution block. This analysis is based on the spatial invariance of fully convolutional neural networks (FCNs). A threshold estimation component is designed to automatically learn an appropriate threshold for each hazy image using convolutional blocks. The threshold distinguishes weak-haze and strong-haze regions and is globally related to the transmission map. The convolution kernel size of the threshold prediction layer matches the size of the transmission map. To reduce the number of parameters, the transmission map is first downsampled by 32 times using average pooling,

$$\mathrm{trans} = \mathrm{Avgpool}(\mathrm{trans}, 32) \tag{3}$$

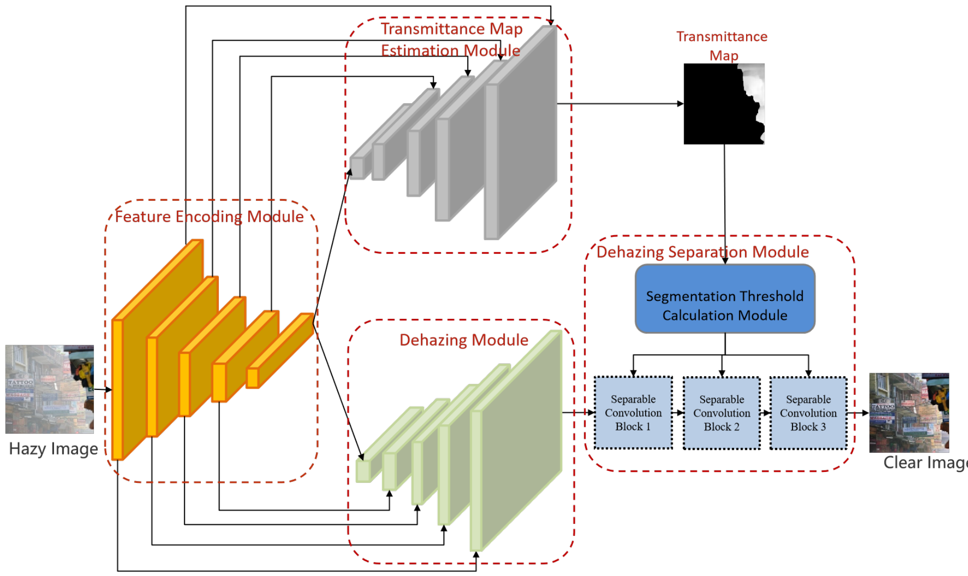

Fig. 1: Dehazing Model Based on Haze Degree Prediction

Next, a convolution layer with a kernel size equal to $1/32$ of the original image size is used to compute the threshold, and then the result is mapped to the range $(0,1)$. The calculation is as follows:

$$\text{Threh} = \frac{1}{1 + e^{-f(\text{trans})}} \tag{4}$$

where trans represents the transmission map, and $f(\text{trans})$ represents the convolution module designed for the transmission map. By comparing the pixel values in the transmission map with the estimated threshold, weak fog region mask $ML(x,y)$ and strong fog region mask $MD(x,y)$ are determined, as shown in equations (5) and (6):

$$ML(x,y) = \begin{cases} 1, & \text{if } \text{trans}(x,y) \geq \text{Threh} \\ 0, & \text{if } \text{trans}(x,y) < \text{Threh} \end{cases} \tag{5}$$

$$MD(x,y) = \begin{cases} 1, & \text{if } \text{trans}(x,y) < \text{Threh} \\ 0, & \text{if } \text{trans}(x,y) \geq \text{Threh} \end{cases} \tag{6}$$

The separable convolution block is based on the spatial invariance of convolutional neural networks. It separates the features $f(x,y)$ based on the weak fog region mask $ML(x,y)$ and the strong fog region mask $MD(x,y)$, resulting in weak fog region features $f_{ML}(x,y)$ and strong fog region features $f_{MD}(x,y)$, as calculated in equations (7) and (8):

$$f_{ML}(x,y) = f(x,y) \times ML(x,y) \tag{7}$$

$$f_{MD}(x,y) = f(x,y) \times MD(x,y) \tag{8}$$

Two parallel convolution blocks process the weak fog region features $f_{ML}(x,y)$ and strong fog region features $f_{MD}(x,y)$. The outputs of the two convolution blocks F_1 and F_2 are added pixel-wise to obtain the complete feature, as shown in the following equation:

$$f_r(x,y) = F_1(f_{ML}(x,y)) + F_2(f_{MD}(x,y)) \tag{9}$$

Finally, a convolutional block further optimizes the features obtained from parallel separation dehazing. The structure of the separation convolution block is shown in Figure 2.

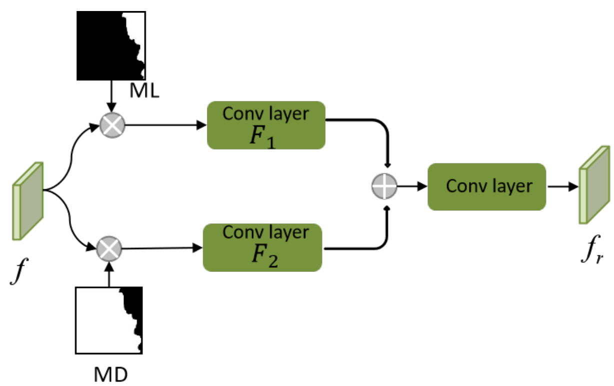

Fig. 2: Structure of the Separation Convolution Block

3.2 Model Structure

The dehazing algorithm model based on haze degree prediction is illustrated in Figure 3. The encoding module comprises five downsampling convolutional blocks, while the transmission map estimation and dehazing modules consist of a bottleneck convolutional layer and five upsampling convolutional blocks. The separation convolution module includes two threshold prediction convolutional blocks and three separation convolution blocks.

Densenet [1] introduced short connections between the inputs and outputs of convolutional layers, effectively reducing gradient vanishing issues caused by deep networks and making training easier. Inspired by Densenet, this paper adopts dense connection layers (Dense_Layer) as the basic convolutional block for Conv_Block, Bottleneck_Conv_Block,

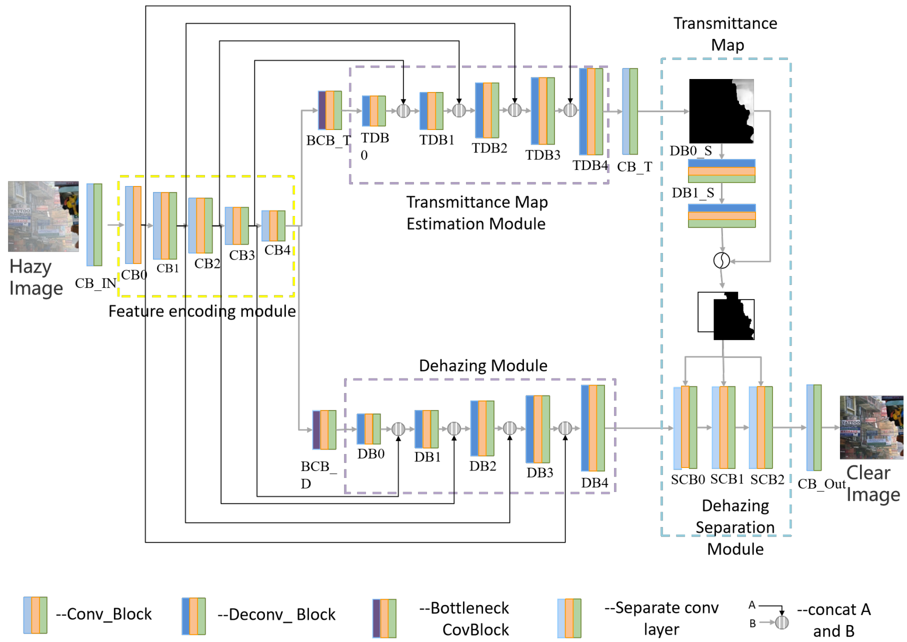

Fig. 3: Network Structure of the Dehazing Model

and Decov_Block. Dense_Layer consists of three sub-convolutional blocks, where the inputs and outputs of each sub-block are channel-wise concatenated. The structures of Conv_Block, Bottleneck_Conv_Block, and Decov_Block are shown in Figure 4, and the basic module is illustrated in Figure 5.

The adaptive threshold calculation module, shown in Figure 6, first downsamples the transmission map by a factor of 32. A convolution layer with a kernel size of 8 learns global features, and the threshold is mapped to the range $(0, 1)$ using a Sigmoid function.

The separation convolution module divides input features using the mask maps. Two convolutional kernels process the divided features separately, and the input and output of each submodule are channel-wise concatenated using the Densenet approach. The specific module structure is illustrated in Figure 7.

3.3 Loss Function

The dehazing network in this paper mainly completes two tasks: one is the estimation of the transmission map, and the other is dehazing. Therefore, the loss function is divided into two parts: the loss function for estimating the transmission map and the dehazing loss function.

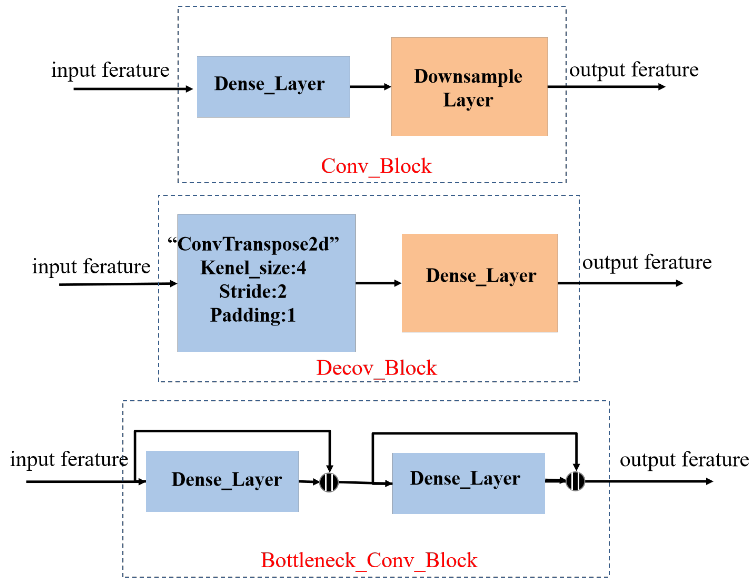

Fig. 4: Submodule Network Structure of the Dehazing Network

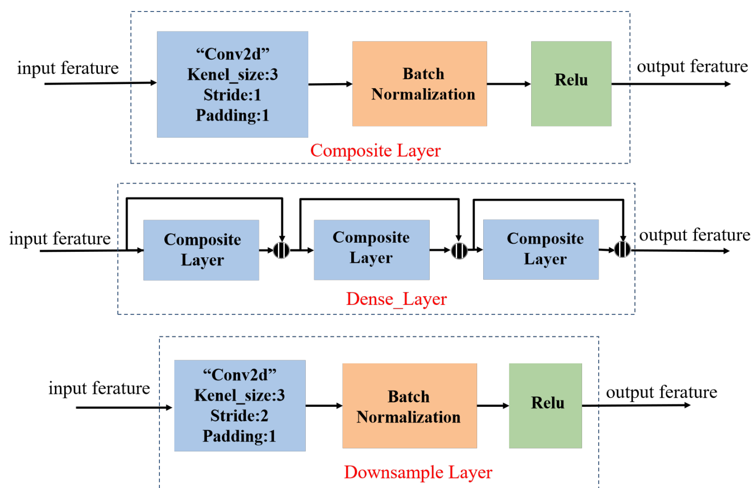

Fig. 5: Basic Module Network Structure

Loss function for estimating the transmission map. The transmission map estimation loss function L_{trans} designed in this paper consists of two parts: a pixel-based error function L_1 loss function $L_{(L1_t)}$ and a gradient loss function L_{grad}. The definition of the

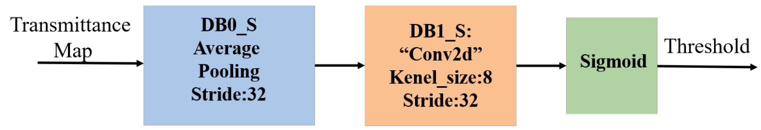

Fig. 6: Adaptive Threshold Calculation Module

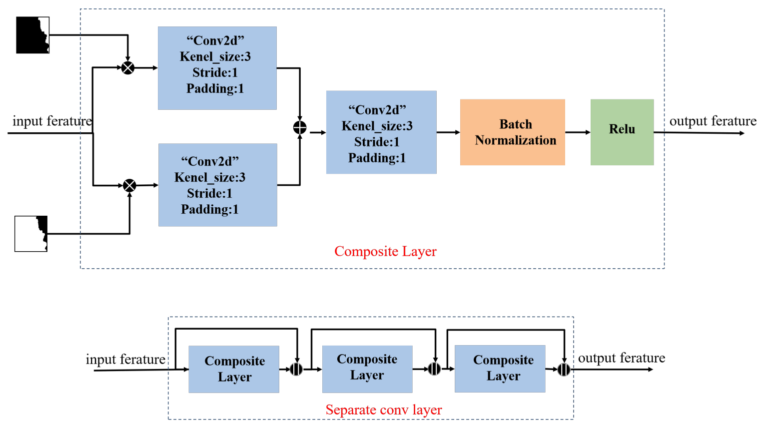

Fig. 7: Separation Convolution Module Network Structure

$L_{(L1_t)}$ loss function is as follows:

$$L_{(L1_t)} = \sum_{w,h,c} \|T(I) - t\|_1 \tag{10}$$

where $T(I)$ represents the transmission map estimation network, I is the input hazy image, t is the target transmission map, and w, h, and c represent the width, height, and channels of the image, respectively.

To optimize the edge information of the transmission map, the gradient loss function L_{grad} is introduced on top of the $L1$ loss function. The calculation is as follows:

$$L_{\text{grad}} = \sum_{w,h} (\|G_x(T(I)) - G_x(t)\|_2 + \|G_y(T(I)) - G_y(t)\|_2) \tag{11}$$

where G_x and G_y represent the horizontal and vertical gradient calculation functions, respectively.

Dehazing loss function. The dehazing loss function is an $L1$-based loss function, defined as follows:

$$L_{\text{dehaze}} = \sum_{w,h,c} \|S - C\|_1 \tag{12}$$

where S is the dehazed clear image, and C is the target clear image.

In summary, the overall loss function for the image dehazing algorithm based on the degree of image haziness in this paper is:

$$L_{\text{sum}} = L_{\text{dehaze}} + L_{\text{trans}} \tag{13}$$

3.4 Evaluation Metrics

Peak Signal-to-Noise Ratio (PSNR) **Peak Signal-to-Noise Ratio (PSNR):** PSNR is commonly used to evaluate the distortion of image reconstruction. The higher the decibel value, the better the image reconstruction quality. In this paper, PSNR is used to calculate the distortion of the dehazed image compared to the real clear image. The formula for PSNR is as follows:

$$\text{PSNR} = 10\log_{10}\left(\frac{p_{\text{peak}}}{\text{MSE}}\right) \tag{14}$$

$$\text{MSE} = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left(p(i,j) - g(i,j)\right)^2 \tag{15}$$

where p_{peak} is the maximum pixel value of the image (255 for grayscale images), MSE is the mean squared error, and $p(i,j)$ and $g(i,j)$ represent the pixel values of the dehazed image and the real clear image, respectively.

Structural Similarity Index (SSIM): SSIM is used to evaluate the structural information between two images. It is a comprehensive metric that considers luminance, contrast, and covariance between two images. The higher the value, the more consistent the structure of the two images. In this paper, SSIM is used to calculate the structural similarity between the real clear image and the dehazed image. The formula for SSIM is as follows:

$$\text{SSIM}(f,g) = \frac{(2\mu_f\mu_g + c_1)(2\sigma_{fg} + c_2)}{(\mu_f^2 + \mu_g^2 + c_1)(\sigma_f^2 + \sigma_g^2 + c_2)} \tag{16}$$

where f and g are the two images being compared, μ_f and σ_f^2 are the mean and variance of image f, μ_g and σ_g^2 are the mean and variance of image g, and σ_{fg} is the covariance between images f and g.

4 Results

4.1 Dataset Introduction

Since there is no real pair of fog images, clear images, and transmission maps available, the widely adopted method is to use a synthesized fog image database from the NYU_depth dataset. However, the scenes in the NYU_depth dataset are all indoor images, while the primary application of dehazing algorithms is outdoor scenes. Therefore, this paper synthesizes a rich scene dehazing image dataset based on the ReDWeb_V1 dataset. The ReDWeb_V1 dataset [10] is a database containing both indoor and outdoor images along with corresponding depth information. Given a clear image and depth map, random values of atmospheric light $(0.8, 1)$ and atmospheric scattering coefficient $(0.4, 1.6)$ are generated to synthesize foggy images based on the atmospheric scattering model. Finally, the sizes of the clear images and the synthesized fog images are standardized to 256×256.

This paper uses the publicly available SOTS dataset from the RESIDE dataset as the test dataset for evaluating algorithm performance. The SOTS dataset contains 500 pairs of outdoor foggy and clear images, which do not overlap with the training dataset.

4.2 Loss Function for Transmission Map Estimation

To demonstrate the effectiveness of the loss function for estimating the transmission map, this paper explores different loss functions for the same network structure: (i) L1 loss, (ii) L1 loss + gradient loss. The experimental results are shown in Figure 8. It can be observed that using both the L1 loss function and the gradient loss function results in smoother edge information in the transmission map.

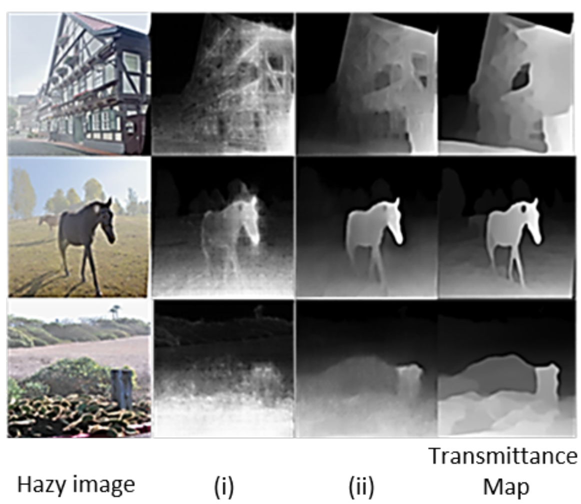

Fig. 8: Effect of Loss Functions on Transmission Map Estimation

4.3 Training Process and Experimental Results

Based on the dehazing network built in Section 2.2.2, experiments were conducted using the synthesized ReDWeb_V1 dataset. After 6000 iterations, the network converged. The loss function curves for the overall network, dehazing loss, and transmission map estimation loss are shown in Figure 9. Overall, the oscillation amplitude decreases and the trend of the loss function continues to decline. After 4000 steps, the decline in the loss function gradually slows down, and the model ultimately converges.

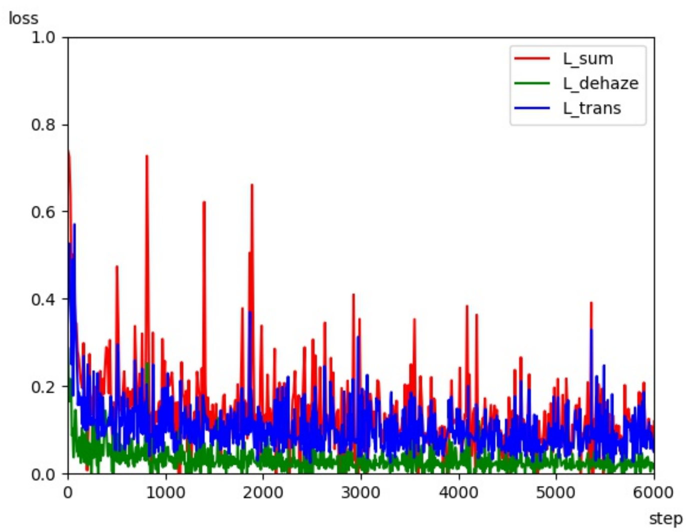

Fig. 9: Dehazing Loss Curve

This paper visualizes the weak and strong fog masks computed by the threshold estimation component, as shown in Figure 10. The first column contains the real-world foggy images, the second column contains the weak fog masks, the third column contains the strong fog masks, and the fourth column shows the estimated transmission maps from the network. The weak fog mask has a pixel value of 1 in areas with less fog, while the strong fog mask is the opposite of the weak fog mask. From the image, we can see that the transmission map reflects the fogging degree of the image, with darker areas corresponding to lower transmission values. Additionally, the mask images obtained from the threshold estimation component effectively segment the image based on fog density, demonstrating the effectiveness of the separation module for parallel dehazing of different fogging levels.

On the outdoor dataset from the public SOTS dataset, this paper compares the proposed algorithm with other well-performing prior knowledge-based dehazing algorithms (DCP,

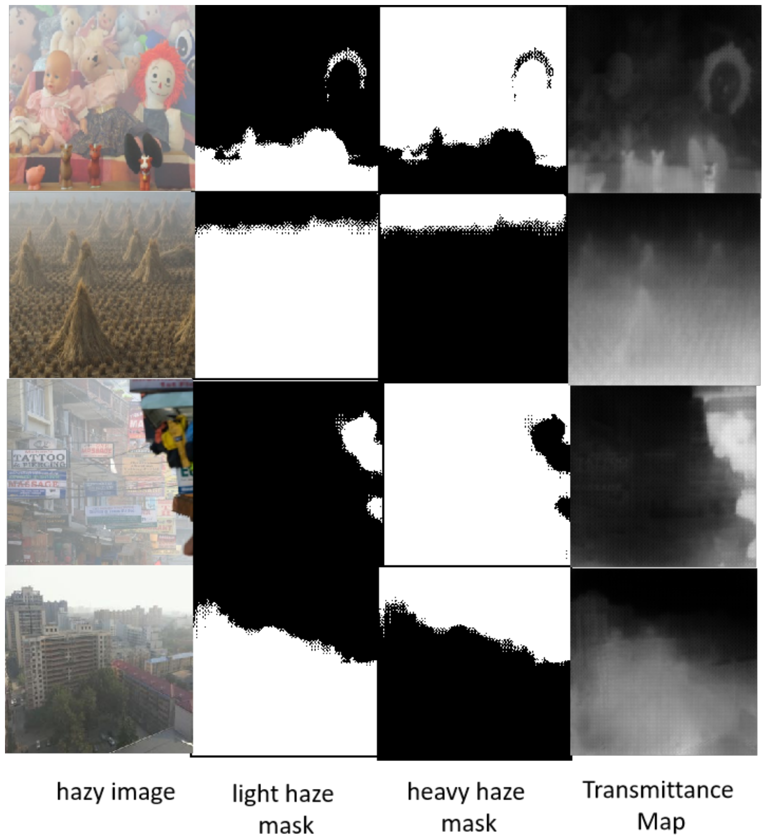

Fig. 10: Weak and Strong Fog Masks Corresponding to Foggy Images

CAP and deep learning-based dehazing algorithms (AOD-net, PFF-net, Dehaze-GAN, DHS-GAN). The comparison results are shown in Table 2.1. Although the proposed algorithm performs slightly worse than DHSGAN in terms of PSNR, it achieves the optimal SSIM value. PSNR calculates the absolute error of each pixel in two images and is an error-sensitive evaluation metric. SSIM, on the other hand, considers the brightness, contrast, and structure of two images and provides higher accuracy in image restoration compared to PSNR. The experiment shows that the proposed algorithm has higher restoration accuracy.

Figure 11 shows the foggy images from six real scenes and their corresponding dehazed results generated by several algorithms. From the first, second, and fifth rows, it can be seen that AOD, Dehaze-GAN, and DHSGAN preserve more haze, while DCP and DHSGAN over-enhance the sky region. In the fourth row, the dehazed results from DCP, AOD, and Dehaze-GAN show dark, information-lost areas in the lower tree sections, while DHSGAN does not

Model	PSNR	SSIM
DCP	19.155	0.899
CAP	18.481	0.773
AOD-net	20.292	0.875
PFF-net	21.252	0.839
Dehaze-GAN	23.595	0.909
DHSGAN	26.612	0.910
Proposed Algorithm	24.827	0.925

Table 1: Dehazing Performance of Different Algorithms

completely remove the fog. In the third row, DCP, AOD, and Dehaze-GAN produce dark buildings (left corner), which contrasts with the result obtained by the proposed algorithm. The last row shows that the proposed dehazing network effectively recovers the true color information, and the sky region's color remains undistorted.

Fig. 11: Dehazing Results for Real-World Foggy Images Using Different Algorithms

5 Conclusion

This paper primarily addresses the issue of varying levels of fog in different regions of a hazy image. A dehazing algorithm based on the fogging degree of the image is proposed. Considering that the transmission map corresponding to a foggy image reflects the fogging degree, the algorithm introduces a transmission map prediction module that runs in parallel with the dehazing module. Based on the obtained transmission map, an adaptive threshold learning module is designed to calculate the segmentation threshold, generating mask images for different levels of fogging. These mask images guide the separation dehazing convolution module to restore a de-fogged image. Through performance evaluation on publicly available datasets and comparisons with dehazing results on real images, the proposed dehazing network achieves higher dehazing accuracy compared to other dehazing algorithms.

References

[1] E. McCartney, "Optics of the atmosphere: scattering by molecules and particles," 1976.

[2] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior.," *IEEE Transactions on Pattern Analysis and Machine Intelligence, Pattern Analysis and Machine Intelligence, IEEE Transactions on, IEEE Trans. Pattern Anal. Mach. Intell*, vol. 33, no. 12, pp. 2341 – 2353, 2011.

[3] Q. Zhu, J. Mai, and L. Shao, "A fast single image haze removal algorithm using color attenuation prior.," *IEEE Transactions on Image Processing, Image Processing, IEEE Transactions on, IEEE Trans. on Image Process*, vol. 24, no. 11, pp. 3522 – 3533, 2015.

[4] D. Berman, S. Avidan, and T. Treibitz, "Non-local image dehazing.," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-December, ((1)Tel Aviv University), pp. 1674–1682 – 1682, 2016.

[5] W. Ren, X. Cao, S. Liu, H. Zhang, J. Pan, and M.-H. Yang, *Single image dehazing via multi-scale convolutional neural networks.*, vol. 9906 LNCS of *Lecture Notes in Computer Science*. (1)Tianjin University: Springer Verlag, 2016.

[6] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "An all-in-one network for dehazing and beyond.," 2017.

[7] K. Mei, A. Jiang, J. Li, and M. Wang, "Progressive feature fusion network for realistic image dehazing.," 2018.

[8] N. Bharath Raj and N. Venketeswaran, "Single image haze removal using a generative adversarial network.," *2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), Wireless Communications Signal Processing and Networking (WiSPNET), 2020 International Conference on*, pp. 37 – 42, 2020.

[9] R. Malav, K. Ayoung, S. R. Sahoo, G. Pandey, C. V. Jawahar, L. Hongdong, G. Mori, and K. Schindler, "Dhsgan: An end to end dehazing network for fog and smoke.," 2019.

[10] K. Xian, C. Shen, Z. Cao, H. Lu, Y. Xiao, R. Li, and Z. Luo, "Monocular relative depth perception with web stereo data supervision.," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Computer Vision and Pattern Recognition (CVPR), 2018 IEEE/CVF Conference on, CVPR*, pp. 311 – 320, 2018.