Adaptive Feature Normalization Network for Single Image Dehazing

Guangrui Hu1, Yueting Yang2, Chuansheng Yang1,3, Chao Wang1,3 and Anhui Tan1,3*

1 School of Information Engineering, Zhejiang Ocean University, Zhoushan 316022, China
2 School of Mathematics and Statistics, Beihua University, Jilin, Jilin, 132013, P.R. China
3 Key Laboratory of Oceanographic Big Data Mining and Application of Zhejiang Province, Zhoushan 316022, China
*Corresponding author’s e-mail: tananhui@zjou.edu.cn

Abstract. In this paper, we design an end-to-end adaptive feature normalization network (AFN-Net) for single image dehazing. In order to fit the function that can recover haze-free images from haze images more effectively, an Adaptive Normalization module (AN) is designed, which can unify the input into the same feature space. A large number of experimental evaluations show that the proposed method is superior to the state-of-the-art (SOTA) method on the benchmark datasets.

1. Introduction
The presence of suspended particles in the air will cause the light to be weakened by interference during the propagation process resulting in blurring and noise in the images obtained by the sensor. The dehazing task is one of the basic but essential computer vision tasks. Most traditional dehazing methods on the strength of atmospheric scattering models [1], and the simplified formula is as follows:

\[ I(x) = J(x)t(x) + A(1 - t(x)), \]  \hspace{1cm} (1)

where \( I(x) \) is the observed hazy image, \( J(x) \) is the corresponding clear haze-free image, \( A \) is the intensity of global atmospheric light, and \( t(x) \) is the medium transmission map.

Early traditional haze-free models often use handcrafted image priors, e.g., [2] used spatial information such as geographic reference digital terrain to help remove the haze of images. [3] introduced the effective dark channel prior to directly evaluate the thickness of the haze in the scene and restore high-quality haze-free images. The rapid development of CNN has made learning-based single image dehazing methods more widely studied. [4] proposed the first end-to-end dehazing model based on CNN, which completes the transformation from haze images to clear images. [5] proposed a feature fusion attention network to restore haze-free images directly.

In this article, we propose an end-to-end AFN-Net to solve the haze-free mission. Inspired by [6] and [7], we utilize a Multistage Boosted module (MS) module based on the Strengthen-Operate-Subtract boosting strategy [8] to restore the haze-free images by using the idea of gradual
enhancement. And the proposed novel AN module is embedded in the MS module to refine the strengthened features.

In summary, the contributions of our article are as follows:

- An AFN-Net is proposed for image dehazing. It achieves better qualitative and quantitative performance than some current SOTA dehazing methods.
- A novel adaptive feature normalization module is designed which can unify the input into the same feature space.

2. Proposed method

In this part, we will explain the overall structure of our network and main modules in detail.

2.1. Adaptive feature normalization network

As shown in Figure 1, hazy images as the input of AFN-Net, it is passed into 3 Multistage Boosted modules (MB), the output features of 3 MB modules are fused through a concatenation. Finally, we use the Feature Attention module (FA) [5] and a residual block [9] to restore the fused information to the haze-free images.

2.2. Multistage boosted module

The MB module which contains the FA and the AN is utilized to restore the haze-free images by using the idea of gradual enhancement. We regard a FA and an AN as a group, and the combination form of each group is expressed by the following formula:

\[ f^{n+1} = g(f^n) + I^n - j^n, \]  

where \( j^n \) is the output of the \( n \)-th group, \( I \) is the input of MS. The function \( f(*) \) represent the processing of the FA module, and the function \( g(*) \) shows processing of the AN module. Each group complies with the Strengthen-Operate-Subtract boosting strategy. Features are first used to extract important information through the FA module. Then add the original feature \( j^0 \) for enhancement. We use the AN module for our steps. Finally, subtract the initial input of each group to get the output of each group.

Figure 1. The adaptive feature normalization network (AFN-Net) architecture.

Figure 2. Architecture of the Adaptive Normalization module.
2.3. Adaptive normalization module
The AN module consists of three parts: a channel attention unit [7], a normalization unit, and a dilation convolution unit, as shown in figure 2. The channel attention unit and pixel attention unit are introduced to emphasize the important factors of different channels to pre-process the feature normalization. The normalization unit can unify the input into the same feature space. We extract and process feature information in different ranges by using multi-scale rate dilation convolution units. The following formula is the main content of the AN module:

\[ x_{\text{norm}} = \gamma(x) \ast \left( \frac{x - \mu(x)}{\sigma(x)} \right) + \beta(x), \]  

where \( \gamma(x) \), \( \mu(x) \), \( \sigma(x) \) and \( \beta(x) \) are the parameters learned from sample \( x \). In general, the AN can align the feature information in the feature space and help the model perform better in the training process.

2.4. Loss function
This article mainly uses \( \ell_1 \) loss to optimize the network. In order to strengthen the network’s dehazing ability, we introduce perceptual loss [10] calculated by MSE loss. The overall loss function of the entire network is as follows:

\[ L = L_{\text{pix}} + L_{\text{per}}, \]

where \( L_{\text{pix}} \) denotes the dehazing pixel loss and \( L_{\text{per}} \) represents the perceptual loss.

Table 1. Achievement table of each method of haze removal in Indoor and Outdoor.

| Method   | Indoor | Outdoor |
|----------|--------|---------|
|          | PSNR   | SSIM    | PSNR   | SSIM    |
| DCP      | 16.62  | 0.8179  | 19.13  | 0.8148  |
| AOD-Net  | 19.06  | 0.8504  | 20.29  | 0.8765  |
| GD-Net   | 32.16  | 0.9836  | 30.86  | 0.9819  |
| FFA-Net  | 36.39  | 0.9886  | 33.57  | 0.9871  |
| Ours     | 37.43  | 0.9901  | 34.54  | 0.9882  |

3. Experiments
3.1. Experiment Setup
The datasets, metrics, and implementation details will be introduced in this part separately.

In this paper, we take Indoor Training Dataset (ITS) and Outdoor Training Dataset (OTS) from RESIDE [11] dataset as the training set of our network. And we test our network on the Synthetic Objective Testing Set (SOTS) of RESIDE which consists of 500 indoor and outdoor images respectively.

The whole network is trained for \( 5 \times 10^5 \) steps on ITS and \( 1 \times 10^6 \) steps on OTS respectively. And the initial learning rate is set as \( 1 \times 10^{-4} \). Additionally, our experiments are all worked by PyTorch 1.6.0 on an NVIDIA GeForce RTX 2080Ti GPU with 11G video memory.

3.2. Comparison with SOTA Methods
In this section, our proposed AFN-Net will be compared with previous SOTA methods including DCP [3], AOD-Net [4], GCA-Net [12], GD-Net [13], and FFA-Net [5] on RESIDE dataset both qualitatively and quantitatively. The average PSNR and SSIM values are reported in table 1. As can be seen from the table, compared with other dehazing methods, our network has the highest score of images dehazing in both indoor and outdoor data sets.
The results of qualitative comparison between these SOTA methods and ours are shown in Figure 3. The first two lines show the results of indoor image dehazing of each method, while the last three lines show the results of outdoor image dehazing. We have marked some dehazing details with red boxes.

**Figure 3.** Qualitative comparison is made on the SOTS data set for six dehazing methods. The first column is hazy image, and the last column is corresponding Ground-truth (GT).

The outcomes by DCP, AOD-Net, and GD-Net remain a little haze. Despite GD-Net, GCA-Net and FFA-Net do well in preserving the colour, they still produce prominent shadows in some areas. And in some places with clear edges, e.g., the selection area in the second row of figure 3, their results only show a blurry edge structure. Notably, the proposed AFN-Net has a better ability of dehazing and can obtain a more realistic dehazing result while preserving the edges.

### 4. Conclusion

In this article, we propose an end-to-end AFN-Net to solve the haze-free mission. we utilize the MS module to restore the haze-free images by using the idea of gradual enhancement. And the proposed novel AN module is embedded in the MS module to refine the strengthened features. It can also be extended to other single image processing tasks, such as deraining, denoising, and super-resolution.

### Acknowledgements

This work was supported by the National Natural Science Foundation of China 12001005; Key Project of Natural Science Foundation Joint Fund of Jilin Province (2020122367JC); Science and Technology Development Project Program of Jilin Province (20190303132SF).

### References

[1] McCartney, E.J. (1976) Optics of the atmosphere: scattering by molecules and particles. New York.

[2] Kopf, J., Neubert, B., Chen, B., Cohen, M., Cohen-Or, D., Deussen, O., ... Lischinski, D. (2008) Deep photo: Model-based photograph enhancement and viewing. ACM transactions on graphics (TOG), 27(5): 1-10.

[3] He, K., Sun, J., & Tang, X. (2010) Single image haze removal using dark channel prior. IEEE transactions on pattern analysis and machine intelligence, 33(12): 2341-2353.

[4] Li, B., Peng, X., Wang, Z., Xu, J., & Feng, D. (2017) Aod-net: All-in-one dehazing network. In Proceedings of the IEEE international conference on computer vision. pp. 4770-4778.
[5] Qin, X., Wang, Z., Bai, Y., Xie, X., & Jia, H. (2020) FFA-Net: Feature fusion attention network for single image dehazing. In Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 34, No. 07, pp. 11908-11915.

[6] Wang, C., Shen, H.Z., Fan, F., Shao, M. W., Yang, C. S., Luo, J. C., & Deng, L. J. (2021) EAA-Net: A novel edge assisted attention network for single image dehazing. Knowledge-Based Systems, 107279.

[7] Hu, J., Shen, L., & Sun, G. (2018) Squeeze-and-excitation networks. In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7132-7141.

[8] Romano, Y., & Elad, M. (2015) Boosting of image denoising algorithms. SIAM Journal on Imaging Sciences, 8(2), 1187-1219.

[9] He, K., Zhang, X., Ren, S., & Sun, J. (2016) Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770-778.

[10] Johnson, J., Alahi, A., Fei-Fei, L. (2016) Perceptual losses for realtime style transfer and super-resolution, in: European conference on computer vision, Springer. pp. 694–711.

[11] Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W., & Wang, Z. (2018) Benchmarking single-image dehazing and beyond. IEEE Transactions on Image Processing, 28(1), 492-505.

[12] Chen, D., He, M., Fan, Q., Liao, J., Zhang, L., Hou, D., … & Hua, G. (2019) Gated context aggregation network for image dehazing and deraining. In 2019 IEEE winter conference on applications of computer vision (WACV). pp. 1375-1383.

[13] Liu, X., Ma, Y., Shi, Z., & Chen, J. (2019) Griddehazenet: Attention-based multi-scale network for image dehazing. In Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 7314-7323.