IEICE TRANSACTIONS

on Information and Systems

This advance publication article will be replaced by the finalized version after proofreading.

IAD-Net: Single-Image Dehazing Network Based on Image Attention

Zheqing ZHANG[†], Hao ZHOU[††*], Chuan LI[†††], *Nonmembers*, *and* Weiwei JIANG[††††], *Member*

SUMMARY   Single-image dehazing is a challenging task in computer vision research. Aiming at the limitations of traditional convolutional neural network representation capabilities and the high computational overhead of the self-attention mechanism in recent years, we proposed image attention and designed a single image dehazing network based on the image attention: IAD-Net. The proposed image attention is a plug-and-play module with the ability of global modeling. IAD-Net is a parallel network structure that combines the global modeling ability of image attention and the local modeling ability of convolution, so that the network can learn global and local features. The proposed network model has excellent feature learning ability and feature expression ability, has low computational overhead, and also improves the detail information of hazy images. Experiments verify the effectiveness of the image attention module and the competitiveness of IAD-Net with state-of-the-art methods.
key words: *image dehazing, self-attention, convolutional neural network, deep learning*

1. Introduction

With the continuous development of social industrialization, air pollution has become more and more serious, and the occurrence frequency of haze in the atmosphere has become more frequent. Atmospheric particles suspended in fog and haze weather absorb and scatter light, resulting in reduced contrast, loss of detail, and color shift in images acquired by imaging equipment, which is not conducive to image recognition and application. The occurrence of fog and haze will not only affect the perception of the pictures taken, but also have a serious impact on the computer vision system whose main work is image processing. The classic image dehazing methods include image enhancement and image restoration.

The dehazing method based on image enhancement is mainly to enhance the contrast and saturation of the details in the haze image, highlighting the interesting parts of human vision. These include histogram equalization [1], Retinex algorithm [2], etc. The methods based on image

restoration mainly use physical models and estimate various parameters to solve clear images. Researchers have proposed various prior knowledge to scientifically estimate parameters in the physical model to improve the effect of dehazing. Among them, the prior with great influence is the method based on dark channel prior (DCP) proposed by He et al [3]. A large number of image dehazing works are improved or expanded based on DCP [4–8]. However, these methods suffer from poor dehazing robustness due to the limitations of physical models and handcrafted priors.

With the rapid development of deep learning, it has been widely used in the field of image dehazing. Early deep dehazing models were based on convolutional neural networks [9–14]. In recent years, with the development of Transformer architecture [15], Transformer-based image dehazing models have been continuously proposed [16–21]. Although the dehazing network based on the Transformer architecture has excellent performance due to its long-distance representation capability, it greatly increases the number of parameters and computational overhead of the network, which is not conducive to deployment in terminal devices with real-time requirements. Therefore, we propose a lightweight dehazing network IAD-Net based on image attention. IAD-Net has a lower number of parameters and computational overhead, and has better dehazing performance. Our contributions can be summarized as:

- We propose a novel attention mechanism, called image attention, which is a plug-and-play module with global modeling capabilities that can better characterize the features of images.

- We design a lightweight depth image dehazing model IAD-Net based on Image Attention. IAD-Net is a parallel lightweight network structure that combines the global modeling ability of image attention and the local modeling ability of convolution, enabling the network to learn global and local features and fuse them. IAD-Net has good feature learning and feature expression capabilities, low computational overhead, and can restore image details while removing haze.

- We conduct experimental comparisons with state-of-the-art methods. Experimental results show that the proposed image attention module can effectively improve the dehazing ability, and IAD-Net is competitive with state-of-the-art image dehazing methods.

[†]The author is with the School of Big Data and Artificial Intelligence, Chizhou University, Chizhou 247000, China.
  [††]The author is with the School of Computer Science and Technology, Anhui University of Technology, Ma'anshan 243032, Anhui, China.
  [†††]The author is with the College of Computer and Information Science, Southwest University, Chongqing 400715, Chongqing, China.
  [††††]The author is with School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, Beijing, China.
  [*]E-mail:haozhou@ahut.edu.cn(Corresponding author)

2. Related Work

The earliest work on the image dehazing algorithm based on deep learning is the trainable end-to-end model DehazeNet proposed by Cai et al. [9], which is used to estimate the transmission map. Li et al. [10] reconstructed the atmospheric scattering model and proposed AODNet. Ren et al. [14] proposed MSCNN to roughly estimate and refine the transmission map. With the development of the deep self-attention structure Transformer [15], many researchers have proposed dehazing networks based on the Transformer structure, such as HyLoG-ViT [16], Transweather [17], DeHamer [18], and DehazeFormer [19]. Although Transformer-based dehazing networks have achieved great dehazing performance, the parameter and computational overhead are large, which limits real-time applications in terminals. Therefore, we propose a novel image attention mechanism and dehazing network IAD-Net, which greatly reduces the number of parameters and computational overhead while ensuring dehazing performance.

3. Proposed Method

3.1 Motivation

The proposed lightweight image dehazing model IAD-Net is an end-to-end structure, which allows the model to directly learn a mapping from a hazy image to a clear image without considering the estimation of parameters based on the physical model. This can well avoid problems such as inaccurate estimation of parameters by various prior knowledge and the limitations of the atmospheric scattering model (ASM). In addition, most existing image dehazing models based on convolutional neural networks continuously increase the depth and width of the network to improve the fitting ability of the network, or improve performance through the Transformer module. However, this will greatly increase the network parameters and make it difficult to train and deploy. Therefore, we propose a novel attention mechanism, called image attention, which has global modeling capabilities and is integrated with traditional convolutional neural network modules to improve performance while reducing model parameters and computational overhead.

3.2 Architecture of IAD-Net

The overall architecture of IAD-Net is shown in Fig. 1, where 1×1 CONV represents a convolution operation of size 1×1, and 3×3 CONV represents a convolution operation of size 3×3. "Image-Attention" represents the proposed image attention module, "Attention" represents the attention calculation module, and "CNN" represents the convolutional neural network module. IAD-Net can be seen as a parallel architecture, that is, the "Image-Attention" module and the "CNN" module are used to extract features from the input hazy images and integrate global and local features to make
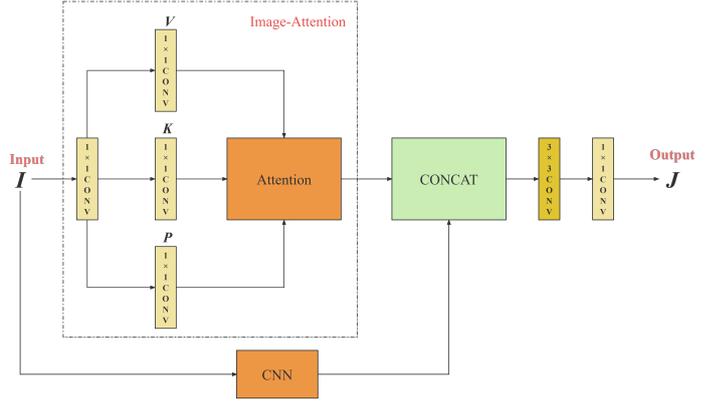
Fig. 1 The architecture of IDA-Net.

Fig. 2 The architecture of CNN module.

the network have better representation ability.

The CNN module in IAD-Net uses the down-sampling and up-sampling modules in the U-Net [22] structure, as shown in Fig. 2. This module can extract multi-scale features, and the number of parameters is very low.

3.3 Module of Image Attention

The proposed image attention module consists of three parallel convolutions and an attention calculation module. First, the input image is extracted through a 1×1 convolution operation, and then different high-level features are extracted through three parallel 1×1 convolutions, named V, K, and P respectively. Then the global features are extracted through the following attention calculation method:

$$Attention(V, K, P) =$$
$$\sum_{j=1}^{N} \sum_{i=1}^{N} (\text{Softmax}(V^j \times \text{Softmax}(P^i \otimes K^j))), \quad (1)$$

The feature map of the input image is padded to a multiple of 16, and then the image is cropped into image blocks of 16×16 size. N represents the number of image blocks, $\times$ represents dot multiplication operation, and $\otimes$ represents matrix multiplication. The proposed image attention calculations are all based on image patches. The calculation is as follows:

(1) Cut the extracted tensor images V, K, and P into N image blocks respectively through parallel convolution operations;

(2) Calculate attention by traversing each image patch and then concat each patch to obtain the attention result of the entire image. The attention calculation method for each image patch is $\text{Softmax}(V^j \times \text{Softmax}(P^i \otimes K^j))$, where i and j represent the patch number, V^j, P^i and K^j represent one of the patches, and Softmax is a common normalization

function in deep learning.

The proposed image attention module combines the advantages of convolution and attention, combining the local feature extraction of convolution with the global feature extraction of attention.

3.4 Loss Function

The proposed IDA-Net is a very lightweight model, and the loss function uses a very simple L1 loss function, as shown below:

$$\zeta = \frac{1}{Num} \sum_{i=1}^{Num} \left\| \boldsymbol{J}^i - \boldsymbol{J}_{GT}^i \right\|_1, \qquad (2)$$

where Num represents the number of images participating in training, $\boldsymbol{J}$ represents the output image obtained by the proposed model, and $\boldsymbol{J}_{GT}^i$ represents the corresponding ground truth (GT) image.

4. Experiments

4.1 Implementation Details

IAD-Net is implemented on devices equipped with GTX 1080Ti GPUs. For synthetic hazy images, IAD-Net is trained on the outdoor subset OTS of the RESIDE dataset [30] and tested on the outdoor test set SOTS-Outdoor. For real-world hazy images, IAD-Net is trained on the O-Haze dataset [23], tested on the NH-Haze dataset [24], and then vice versa.

4.2 Comparison with State-of-the-Art Methods

To verify the performance of IAD-Net, 8 state-of-the-art dehazing methods were selected for qualitative and quantitative experimental comparisons, including: DCP [3], DehazeNet [9], AODNet [10], GridDehazeNet [25], FFANet [26], RefineDNet [27], PSD [28], and D4 [29]. We use the codes and models disclosed by these methods to conduct experimental comparisons on the SOTS-Outdoor, O-Haze, and NH-Haze datasets. The comparison of visual dehazing effects is shown in Fig. 3. The images restored by DCP and AODNet are darker. In addition, the images restored by DCP, AODNet, RefineDNet, and PSD have certain distortions in the sky area. DehazeNet, GridDehazeNet, FFANet, and D4 make it difficult to eliminate the influence of haze in real-world hazy images, but our method has an excellent restoration effect on both synthetic hazy images and real-world hazy images and is closest to the ground truth image.

In addition, quantitative experimental comparisons were conducted on all images of these three datasets, and the average evaluation indexes PSNR and SSIM scores were calculated. As shown in Table 1, our method achieved the highest score, which strongly illustrates the excellent performance of the proposed method.

Table 1 Quantitative experimental comparison results of dehazed images. **Bold** numbers indicate the highest indicator score, and underlined numbers indicate the next highest score.

Methods	SOTS-Outdoor		NH-Haze		O-Haze	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DCP [3]	14.80	0.8029	11.51	0.4181	13.41	0.5011
DehazeNet [9]	18.69	0.7426	11.75	0.4431	14.49	0.6006
AODNet [10]	19.64	0.8920	11.81	0.4226	15.01	0.5413
GridDehazeNet [25]	30.34	0.9797	10.43	0.3864	15.26	0.6325
FFANet [26]	30.12	0.9721	11.81	0.4688	14.12	0.5910
RefineDNet [27]	20.77	0.9093	12.65	0.5126	17.21	0.6629
PSD [28]	15.31	0.7621	10.21	0.5254	10.39	0.6015
D4 [29]	25.06	0.9396	12.57	0.5041	15.67	0.6518
IAD-Net (Ours)	**31.41**	**0.9799**	**17.59**	**0.6544**	**18.63**	**0.6877**

4.3 Ablation Study

This section mainly studies the impact of IAD-Net's Image-Attention module and CNN module on network performance through ablation experiments. We removed these two modules respectively, retrained on the ITS dataset, and then tested on the SOTS-Indoor dataset. The results are shown in Table 2. It can be found that adding each module can improve the PSNR and SSIM scores, which illustrates the effectiveness of these two modules.

Table 2 Ablation study of the network. Where *w/o* means to remove the module.

Metrics	SOTS-Indoor	
	PSNR	SSIM
w/o Image-Attention module	23.19	0.8947
w/o CNN module	25.68	0.9030
Ours Full Model	**32.01**	**0.9754**

5. Conclusions

In this letter, we propose a novel lightweight deep dehazing model IAD-Net, which contains a novel image attention module (Image-Attention) and is integrated with the traditional convolutional neural network. IAD-Net has the local feature extraction capabilities of traditional convolution and the long-distance modeling capability of the attention mechanism. It can effectively extract local and global features of images and has good representation capabilities. Experimental results show that IAD-Net has excellent performance on both synthetic hazy images and real-world hazy images. However, our model also has some shortcomings, such as the restored images have slight color distortion. This may be due to the failure of our model to extract color information well. This is a direction for further research in the future.
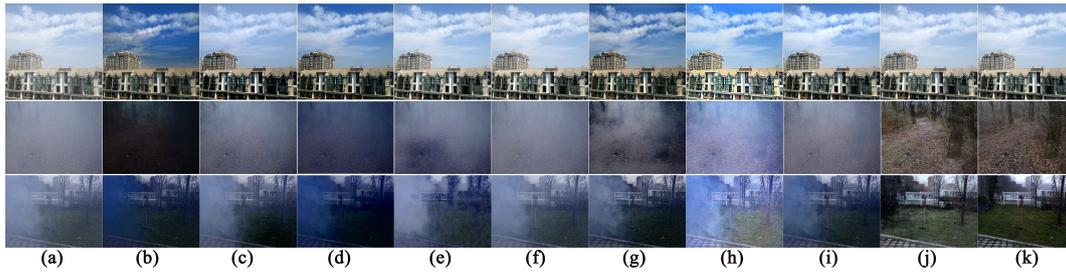
Fig. 3 Dehazing results of different methods. The first image is a synthetic image from SOTS-Outdoor, and the second and third images are real-world images from O-Haze and NH-Haze, respectively. (a) Hazy input. (b) DCP. (c) DehazeNet. (d) AODNet. (e) GridDehazeNet. (f) FFANet. (g) RefineDNet. (h) PSD. (i) D4. (j) Our method. (k) Ground truth.

References

[1] W. Kim, J. You, and J. Jeong. Contrast enhancement using histogram equalization based on logarithmic mapping. *Optical Engineering*, 51(6):067002, 2012.

[2] H. Li, W. H. Xie, X. G. Wang, S. S. Liu, Y. Y. Gai, and L. Yang. Gpu implementation of multi-scale retinex image enhancement algorithm. In *2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)*, pages 1–5, 2016.

[3] K. M. He, J. Sun, and X. O. Tang. Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33:2341–2353, 2011.

[4] Z. Wei, G. Zhu, X. Liang, W. Liu. An image fusion dehazing algorithm based on dark channel prior and retinex. *International Journal of Computational Science and Engineering*, 23:115–123, 2020.

[5] Y. Xiaoping, C. Yanyu, S. Hui. Improved Image Dehazing Algorithm Based on Haze-line and Dark Channel Prior. *LASER & OPTOELECTRONICS PROGRESS*, 59, 2022.

[6] Y. Wang, T.Z. Huang, X.L. Zhao, L.J. Deng, T.Y.Ji. A convex single image dehazing model via sparse dark channel prior. *Applied Mathematics and Computation*, 375: 125085,2020.

[7] Z. Lu, B. Long, S. Yang. Saturation Based Iterative Approach for Single Image Dehazing. *IEEE Signal Processing Letters*, 27: 665–669, 2020.

[8] Y. Liu, Z. S. Yan, J. G. Tan, Y. C. Li. Multi-Purpose Oriented Single Nighttime Image Haze Removal Based on Unified Variational Retinex Model. *IEEE Transactions on Circuits and Systems for Video Technology*, 33: 1643–1657, 2023.

[9] B. Cai, X. M. Xu, K. Jia, C. M. Qing, and D. C. Tao. Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 25:5187–5198, 2016.

[10] B. Li, X. Peng, Z. Y. Wang, and D. Xu, J. Z.and Feng.

Aod-net: All-in-one dehazing network. In *Proceedings IEEE International Conference on Computer Vision*, pages 4780–4788, 2017.

[11] W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W. Liu, M.H. Yang. Gated fusion network for single image dehazing. In *Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 3253–3261.

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

[13] X. Qin, Z. Wang, Y. Bai, X. Xie, H. Jia. FFA-Net: Feature fusion attention network for single image dehazing. In *Proceedings of the Proceedings of the AAAI conference on artificial intelligence*, 2020, Vol. 34, pp. 11908–11915.

[14] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, M.H. Yang. Single image dehazing via multi-scale convolutional neural networks. In *Proceedings of the European conference on computer vision*, Springer, 2016, pp. 154–169.

[15] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, Polosukhin I. Attention is all you need. *Advances in Neural Information Processing Systems*, 2017, Vol. 30.

[16] D. Zhao, J. Li, H. Li, L. Xu. Hybrid local-global transformer for image dehazing. *ArXiv Preprint ArXiv:2109.07100*, 2021, p.

[17] J. M. J. Valanarasu, R. Yasarla, V. M. Patel. Transweather: Transformer-based restoration of images degraded by adverse weather conditions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 2353–2363.

[18] C. L. Guo, Q. X. Yan, S. Anwar, R. M. Cong, W. Q. Ren, C. Y. Li. Image Dehazing Transformer with Transmission-Aware 3D Position Embedding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 5812–5820.

[19] Y. D. Song, Z. Q. He, H. Qian, X. Du. Vision transformers for single image dehazing. *IEEE Transactions on Image Processing*, 2023, Vol. 32, pp. 1927–1941.

[20] H. Zhou, Z. K. Chen, Y. Liu, Y. P. Sheng, W. Q. Ren, H. L. Xiong. Physical-priors-guided Dehaze-Former. *Knowledge-Based Systems*, 2023, Vol. 266, pp. 110410.

[21] Y. Liu, Z. S. Yan, S. X. Chen, T. Ye, W. Q. Ren, E, Chen. J. G. Tan, Y. C. Li. Nighthazeformer: Single nighttime haze removal using prior query transformer. In *Proceedings of the 31st ACM International Conference on Multimedia*, 2023, pp. 4119–4128.

[22] S. W. Zhang, C. L. Zhang. Modified U-Net for plant diseased leaf image segmentation. *Computers and Electronics in Agriculture*, 2023, Vol. 204, pp. 107511.

[23] C. O. Ancuti, C. Ancuti, R. Timofte, C.D. Vleeschouwer. O-haze: A dehazing benchmark with real hazy and haze-free outdoor images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 754–762.

[24] C. O. Ancuti, C. Ancuti, R. Timofte. NH-HAZE: An image dehazing benchmark with non-homogeneous hazy and haze-free images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 444–445.

[25] X. H. Liu, Y. R. Ma, Z. H. Shi, J. Chen. Griddehazenet: Attention-based multi-scale network for image dehazing. In *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 7314–7323.

[26] X. Qin, Z. Wang, Y. Bai, X. Xie, H. Jia. FFA-Net: Feature fusion attention network for single image dehazing. In *Proceedings of the AAAI conference on artificial intelligence*, 2020, Vol. 34, pp. 11908–11915.

[27] S. Zhao, L. Zhang, Y. Shen, Y. Zhou. RefineDNet: A weakly supervised refinement framework for single image dehazing. *IEEE Transactions on Image Processing*, 2021, Vol. 30, pp. 3391–3404.

[28] Z. Chen, Y. Wang, Y. Yang, D. Liu. PSD: Principled synthetic-to-real dehazing guided by physical priors. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 7180–7189.

[29] Y. Yang, C. Wang, R. Liu, L. Zhang, X. Guo, D. Tao. Self-augmented unpaired image dehazing via density and depth decomposition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 2037–2046.

[30] B. Y. Li, W. Q. Ren, D. P. Fu, D. C. Tao, D. Feng, W. J. Zeng, Z. Y. Wang. Reside: A benchmark for single image dehazing. *IEEE Transactions on Image Processing*, 2019, Vol. 28, pp. 492–505.