



CP-Net: Channel Attention and Pixel Attention Network for Single Image Dehazing

Shunan Gao, Jinghua Zhu^(✉), and Yan Yang

School of Computer Science and Technology, Heilongjiang University,
Harbin 150080, China
zhujinghua@hlju.edu.cn

Abstract. An end-to-end channel attention and pixel attention network (CP-Net) is proposed to produce dehazed image directly in the paper. The CP-Net structure contains three critical components. Firstly, the double attention (DA) module consisting of channel attention (CA) and pixel attention (PA). Different channel features contain different levels of important information, and CA can give more weight to relevant information, so the network can learn more useful information. Meanwhile, haze is unevenly distributed on different pixels, and PA is able to filter out haze with varying weights for different pixels. It sums the outputs of the two attention modules to improve further feature representation which contributes to better dehazing result. Secondly, local residual learning and DA module constitute another important component, namely basic block structure. Local residual learning can transfer the feature information in the shallow part of the network to the deep part of the network through multiple local residual connections and enhance the expressive ability of CP-Net. Thirdly, CP-Net mainly uses its core component, DA module, to automatically assign different weights to different features to achieve satisfactory dehazing effect. The experiment results on synthetic datasets and real hazy images indicate that many state-of-the-art single image dehazing methods have been surpassed by the CP-Net both quantitatively and qualitatively.

Keywords: Image dehazing · Channel attention and pixel attention · Residual learning

1 Introduction

In the past 20 years, the problem of image dehazing has aroused wide attention in the computer vision field. Haze is a common atmospheric phenomenon caused by small floating particles such as dust and smoke in the air. These floating particles greatly absorb and scatter light, resulting in reduced image quality. Under the influence of haze, many practical applications such as video surveillance, remote sensing, and autonomous driving are vulnerable to threat, and advanced computer vision jobs such as segmentation [23–25] and object detection [11, 21, 22] are difficult to complete. Therefore, image dehazing has become

an increasingly important technique and its purpose is to restore a hazy image to a haze-free image (see Fig. 1). Most of the successful approaches depend on the physical scattering model [3–5], which can be expressed by the following formula

$$I(z) = J(z)t(z) + A(1 - t(z)) \quad (1)$$

Where I is a hazy image, J is a haze-free image, t is the transmission map and A is the global atmosphere light.

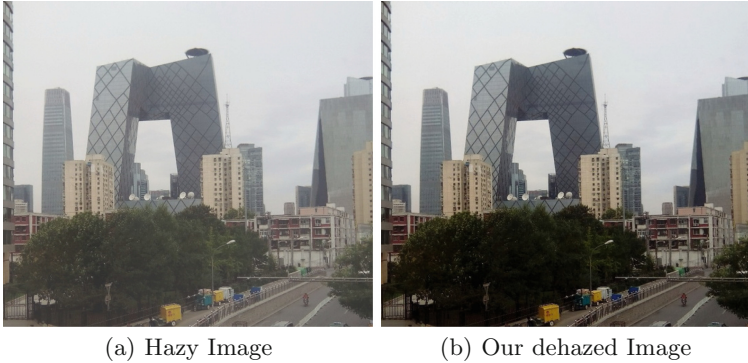


Fig. 1. An example of image dehazing

When the atmosphere is uniform, the transmission map t can be expressed as

$$t(z) = e^{-\beta d(z)} \quad (2)$$

where β is the atmosphere scattering parameter and d is the scene depth.

Equation 1 can also be transformed into the following form

$$J(z) = \frac{I(z) - A}{t(z)} + A \quad (3)$$

We know from Eq. 3 that since A and t have an infinite number of solutions, it is a pathological problem to use the atmospheric scattering model to dehaze. If A and t can be appropriately evaluated by leveraging the captured hazy image, a clear dehazing image we can get. However, it is often challenging to complete that.

Many early dehazing methods, such as [8, 18, 26, 27] were based on atmospheric scattering models. [9, 20] discovered the effective dark channel prior (DCP). In a haze-free image, every local area is likely to have shadows, or something of pure color, or something of black. Therefore, it is very likely that each local area will have at least one color channel with a low value. This statistical law is called Dark Channel Prior. However, the dark channel prior is not very suitable for images with the sky. DCP points out some pixels always have at

least one color channel with a very low value in most local areas. This also indirectly proves that different channel characteristics contain different degrees of important information. The representation of network will be greatly limited if all features are treated equally.

Convolutional neural network (CNN) are widely used in computer vision. Many methods based on convolutional neural networks have excellent results in the field of image processing. One of the advantages of deep CNN model compared to traditional image processing algorithms is that it avoids the complex pre-processing of the image, especially manual participation in the image pre-processing process. Up to now, CNN has been widely used in various image-related applications. In the field of image dehazing, almost all methods are based on CNN. For example, DCP, AOD-Net [2], MSCNN [9], EPDN [31]. These dehazing methods have achieved very remarkable results.

The attention mechanism [14–16] is often used in convolutional neural networks, because it can improve the performance of the network remarkably. Inspired by the work [12], a new double attention (DA) module is proposed. Our DA module combines channel attention and pixel attention. DA can give different channel features and pixel features different weights. For example, important feature will be given greater weight, less important feature will be given less weight.

The deeper (complex, more parameters) CNN network is, the more expressive it is. [10] proposed ResNet which is conducive to training deep models based on convolutional neural networks. We incorporate the attention mechanism and the skip connection into DA module. CP-Net utilizes multiple local residual connections to not only transmit the information in the shallow part of the network to the deep part of the network as completely as possible, but also deepen the depth of the network and improve the network performance.

Many dehazing methods nowadays utilize peak signal to noise ratio (PSNR) and structure similarity (SSIM) indicators to measure the quality of dehazed image recovery. For human subjective evaluation, we also provide a large number of network outputs from corrupted inputs. Experimental results demonstrate that our network exceeds the previous state-of-the-art methods in terms of both the PSNR and SSIM metrics and qualitative comparisons. Not only that, we also made an ablation analysis to prove the effectiveness of our DA module.

The main contributions of our work can include the following:

1. We design a fully end-to-end single image dehazing algorithm CP-Net, which can directly output dehazing image without relying on the atmosphere scattering model in Eq. 1. It achieves state-of-the-art performance on both synthetic and real hazy images;
2. We combine the channel attention and pixel attention mechanism to design a new double attention (DA) module. DA module can focus more attention on important channel information and pixel information. The sum of the outputs of two attention module can further improve the capacity of CP-Net;
3. We combine double attention (DA) module and local residual learning to design a new basic block. Local residual learning can transfer the feature

information in the shallow part of the network to the deep part of the network through multiple local residual connections, DA module can give different channel features and pixel features different weights;

4. Our CP-Net contains multiple basic block connections, which not only reduces the loss of information during the flow, but also increases the depth of the network. In addition, CP-Net can also automatically learn different weights for different features.

2 Related Works

Most previous dehazing methods rely on Eq. 1. As we mentioned above, the atmosphere scattering model is ill-conditioned because global atmospheric light and transmission map can not be accurately estimated. No matter which method is used, we can not escape obtaining accurate transmission map and global atmospheric light. Traditional methods can only use different image statistical priors as constraints to minimize the information loss caused by corruption procedure. Modern methods can only use a large number of constraints to learn useful information in image continuously, but the structural design of dehazing network limits the performance.

The dehazing method introduced in [6] makes use of the Dark Channel Prior (DCP), which estimates the transmission map. However, this prior is proved to be unreliable when the scene objects are similar to the atmospheric light. Color attenuation prior was proposed by [7]. It points out that there is a linear relationship between the depth, brightness, and saturation of the hazy image, which can be used to form a function formula. The measure of calculating the optical transmission of foggy scene based on a single input image is proposed by [28], which assumed a premise that the surface shadow and the transfer function are statistically independent. [19] proposed a method (NLD) for estimating transmission map based on global priors to recover the clean image, and the algorithm assumes that the color of pixel points in a haze-free image can be clustered into hundreds of compact clusters in RGB space. Because the prior is based on the assumption under ideal conditions, the prior-based dehazing methods will fail in certain scenarios, such as bad natural weather.

In recent years, deep learning technology has been greatly developed. The emergence of large synthetic datasets [6] have solved the problem of data scarcity, which has also directly promoted the widespread development of data-driven image dehazing algorithms. Although these algorithms reduce the reliance on handmade priors, they still rely on the traditional strategies mentioned above. For example, DehazeNet [1] is an end-to-end system that directly learns and estimates the mapping relationship between a hazy image and its transmission map. [9] leverages a Multi-Scale CNN (MSCNN) that can estimate the transmission very well.

By transforming Eq. 1, the AOD-Net [2] no longer needs to estimate transmission map and atmospheric light. [13] propose a hazy image restoration method based on threshold fusion network (GFN) which consists of an encoding and

decoding network. EPDN [31] simplifies the image dehazing problem to the image-to-image translation problem, which is embedded by the generative adversarial network (GAN) and does not depend on the physical scattering model.

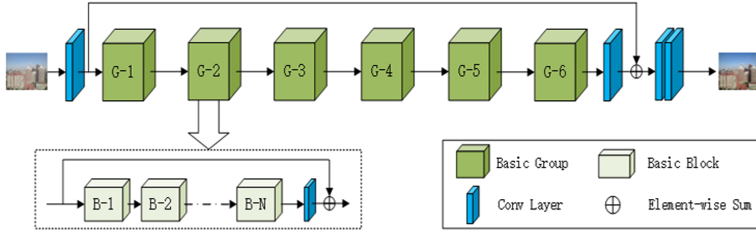


Fig. 2. Channel Attention and Pixel Attention Network (CP-Net) architecture

3 Channel Attention and Pixel Attention Network (CP-Net)

The detailed structure of CP-Net is described in Fig. 2. The input of CP-Net is a hazy image, which first goes through a convolutional layer, then is sent to six Group Architectures with multiple long skip connections and a convolutional layer. The output of convolutional layer is then fused with the output of shallow feature extraction part via element-wise addition. Finally, the features will flow into the reconstruction part, and then a dehazed image will be obtained.

In addition, local residual learning and B Basic Block structures constitute a Group Structure; Double Attention (DA) module and the skip connection constitute a Basic Block. Channel Attention and Pixel Attention constitute DA module.

We will introduce DA module and Basic Block structure in detail in Sect. 3.1 and 3.2 respectively. Finally, we will introduce Group Architecture and Global Residual Learning in detail in Sect. 3.3.

3.1 Double Attention (DA)

Because the distribution of haze on the image is not uniform, the network we designed can deal with different channel features and pixel features differently. Double Attention (see Fig. 3) is comprised of pixel attention and channel attention. DA can help our CP-Net assign different weights to each Channel of input features, extract more critical and important information, and make more accurate judgments of CP-Net. Next, we will elaborate on how CP-Net assigns weights to each channel feature and pixel feature.

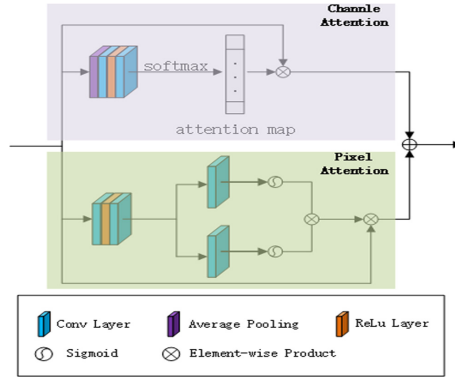


Fig. 3. Double Attention module

Channel Attention (CA). The main task of channel attention is to assign different weights to different channel features. By using the global average pooling, we transform the global spatial information on the channel into a channel descriptor.

$$g_c = G_p(F) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W F_c(i, j) \tag{4}$$

Where G_p is the global pool function, $F_c(i, j)$ represents the pixel value of the position (i, j) on the c -th channel F_c , and the feature map with the shape of $C \times H \times W$ turns into the attention map of $C \times 1 \times 1$. The features are processed by two convolutional layers, a ReLu activation function and a latter softmax function, and then the attention map is obtained, that is, the weights of different channels.

$$CA_c = S(Conv(\delta(Conv(g_c)))) \tag{5}$$

Where the S and δ represent the softmax and ReLu functions, respectively.

By merging the weights of the channel CA_c and the input F by element-wise multiplication, we can get the output of the CA module.

$$F_c = CA_c \otimes F \tag{6}$$

Pixel Attention (PA). Since the haze is distributed differently in different pixels of hazy image, the pixel attention (PA) module we proposed can learn the informative contextual feature of each pixel. Each pixel in the hazy image is treated differently by PA module which can pay more attention to those critical pixels.

Firstly, the input F are processed by two convolutional layers and a ReLu activation function.

$$F' = Conv(\delta(Conv(F))) \tag{7}$$

Then the feature map F' passes through a convolution layer and a sigmoid activation function, respectively.

$$F_1 = \sigma(Conv(F')) \tag{8}$$

$$F_2 = \sigma(Conv(F')) \tag{9}$$

Where the σ represents the sigmoid function. The shape of the feature map F' , F_1 and F_2 remain unchanged as $C \times H \times W$.

In order to get pixel attention weight map, we merge F_1 and F_2 by element-wise multiplication.

$$PA = F_1 \otimes F_2 \tag{10}$$

At the end, we get the output of PA module by fusing PA and F by element-wise multiplication.

$$F_p = PA \otimes F \tag{11}$$

As shown in Fig. 3, we finally merge F_c and F_p by element-wise sum to further improve the performance of the network. The final output of DA module is F^* .

$$F^* = F_c + F_p \tag{12}$$

3.2 Basic Block Structure

As is shown in Fig. 4, the basic block structure is composed of double attention (DA) module and local residual learning. Local residual learning can transfer the feature information in the shallow part of the network to the deep part of the network through multiple local residual connections, DA module can give different channel features and pixel features different weights.

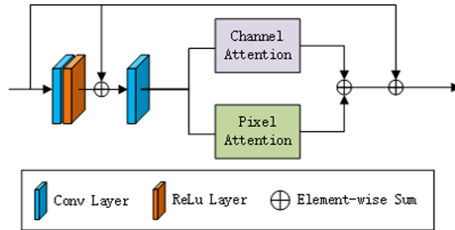


Fig. 4. Basic block structure

3.3 Group Architecture and Global Residual Learning

Our CP-Net contains six Group Architectures. A Group Architecture contains a skip connection and B basic blocks. Our B basic blocks are connected in sequence, followed by a convolutional layer. We fuse the input of Group Architecture and the output of the last convolutional layer by element-wise sum, which not only helps to reduce the loss of information in the flow process, but also helps to deepen the depth of the network. We added three additional convolutional layers and a global learning module after the last Group Architecture. Combining the features of the shallow part of the CP-Net and the features of the deep part through element-wise sum can significantly improve the dehazing effect of our network.

3.4 Loss Function

To train the proposed CP-Net, L1 loss, perceptual loss and SSIM loss are employed.

L1 Loss. Through L2 loss is widely used in many image dehazing networks, [32] proved that L1 loss can achieve higher PSNR and SSIM than L2 loss. Given an input hazy image I , the output of CP-Net is $CP(I)$ and the ground truth is J . Then the L1 loss over N samples can be written as

$$L_1 = \sum_{i=1}^N \|CP(I_i) - J_i\|_1 \quad (13)$$

Perceptual Loss. The perceptual loss leverages multi-scale features extracted from a pre-trained deep neural network to quantify the visual difference between the estimated image and the ground truth. In this paper, we adopt the VGG16 [30] pre-trained on ImageNet [29] and the three stages (ReLU1-2, ReLU2-2 and ReLU3-3). The perceptual loss is defined as

$$L_p = \sum_{j=1}^3 \frac{1}{C_j \times H_j \times W_j} \|\phi_j(CP(I)) - \phi_j(J)\|_2^2 \quad (14)$$

where $\phi_j(CP(I))$ ($\phi_j(J)$), $j = 1, 2, 3$, denote the aforementioned three VGG16 feature maps associated with the dehazed image $CP(I)$ and the ground truth J , and C_j , H_j and W_j specify the dimension of $\phi_j(CP(I))$ ($\phi_j(J)$), $j = 1, 2, 3$.

SSIM Loss. SSIM is proposed to measure the structural similarity between two images. It can be written as

$$SSIM(CP(I), J) = \frac{2\mu_{CP(I)}\mu_J + C_1}{\mu_{CP(I)}^2 + \mu_J^2 + C_1} \cdot \frac{2\sigma_{CP(I)J} + C_2}{\sigma_{CP(I)}^2 + \sigma_J^2 + C_2} \quad (15)$$

Where μ_x, σ_x^2 are the average value and the variance of x , respectively. σ_{xy} is the covariance of x and y . C_1, C_2 are constants used to maintain stability. SSIM ranges from 0 to 1. SSIM Loss can be expressed by the following formula

$$L_s = 1 - SSIM(CP(I), J) \quad (16)$$

Total Loss. We take the sum of the L1 loss, the perceptual loss and the SSIM loss as the total loss L

$$L = L_1 + \alpha L_p + \beta L_s \quad (17)$$

Where α and β are positive weights. We set α and β to be 0.04 and 0.5, respectively.

3.5 Detailed Implementation of CP-Net

In this section, we specify the implementation details of our proposed CP-Net. We set up 6 Group Structures. Each Group Structure contains B=14 Basic Blocks. The convolution layers kernel size of Channel Attention and Pixel Attention is set to 1×1 , but other convolution layers kernel size is 3×3 . Every Group Structure and every Basic Block Structure output 64 features which keep size fixed.

4 Experiments

4.1 Datasets and Metrics

During experiment, we used a synthetic dataset RESIDE [11] containing indoor and outdoor scenes. The indoor dataset contains 28850 hazy images and 2885 clear images for training. These hazy images are generated by corresponding clean images. The outdoor dataset contains 31430 hazy images and 898 clear images. The scatter parameters range from 0.04 to 0.2; the global atmosphere light changes from 0.8 to 1.0. Synthetic Objective Testing Set (SOTS) is used as the test dataset which contains 500 outdoor and 500 indoor images. We compare PSNR, SSIM and visualized dehazing results with previous state-of-the-art dehazing methods on our test dataset. In addition, we also conducted comparison experiments on Realistic Hazy Images.

4.2 Training Settings

In order to get a fine dehazing effect, random rotation and horizontal flip strategies are adopted to augment the training dataset. Two patches with a size of 240×240 are randomly cropped on the hazy image and its corresponding haze-free image as the inputs of CP-Net. CP-Net is trained for $1 \times 10^6, 1 \times 10^5$ steps on indoor and outdoor datasets, respectively. CP-Net is optimized by Adam Optimizer, where the number of β_1 and β_2 is set up 0.9 and 0.999, respectively.

We set the initial learning rate as 1×10^{-4} . We leverage cosine annealing strategy [17] to adjust the learning rate from the initial value to 0. We define T as the total number of training steps, and η as the initial learning rate. The learning rate η_t is shown below at step t .

$$\eta_t = \frac{1}{2} \left(1 + \cos \left(\frac{t\pi}{T} \right) \right) \eta \quad (18)$$

All experiments are carried out with a Tesla V100 GPU.

4.3 Results on RESIDE Dataset

We compare our CP-Net both quantitatively and qualitatively with the previous state-of-the-art dehazing methods that include DCP, NLD, AOD-Net and EPDN. Table 1 shows the quantitative comparisons of our CP-Net and other networks in terms of PSNR and SSIM. It is clear that our experimental results are superior to previous advanced dehazing methods in terms of PSNR and SSIM. Not only that, but we also made a comparison of visualization in Fig. 5 for qualitative comparisons.

Table 1. Quantitative comparisons on SOTS for different methods.

Methods	Indoor		Outdoor	
	PSNR	SSIM	PSNR	SSIM
DCP	16.52	0.7433	17.48	0.7081
NLD	19.73	0.8043	17.72	0.8413
AOD-Net	21.15	0.8654	22.98	0.8982
EPDN	26.89	0.9401	22.05	0.9186
Ours	31.72	0.9858	30.47	0.9815

As is shown in Fig. 5, the top two lines are the results of indoor comparison, and the bottom two lines are the results of outdoor comparison. We can clearly discover that the dehazing images produced by DCP are very different from the real clear images (GT) in color, and the image details are seriously lost. This is because DCP uses prior assumptions. Images recovered by The NLD have a lot of black spots and the sky is highlighted. The dehazed images produced by AOD-Net have a little color distortion and some residual haze. However, our network can well preserve the true details of the images whether it is processing indoor or outdoor scenes. We can hardly see residual haze on our restored images.

We also evaluate the results on real hazy images and observe that although all models are trained with outdoor synthetic dataset RESIDE, our model can still produce better dehazed images. To a certain extent, our network can effectively remove haze, while maximizing the retention of image details. However, the images recovered from other networks not only have a lot of haze, but also are not as good as our network in terms of color fidelity and image detail.

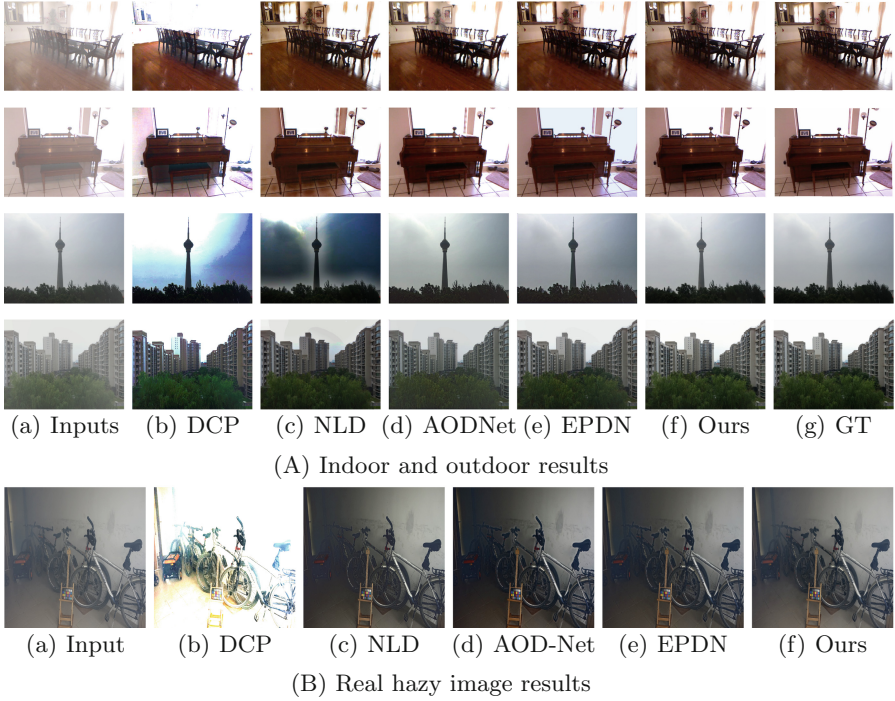


Fig. 5. Qualitative comparisons on SOTS and Realistic Hazy Images testset

5 Ablation Analysis

In order to prove the rationality of our proposed CP-Net structure, we also designed two other networks defined as C-Net and P-Net. The only difference with CP-Net is that C-Net and P-Net contain only Channel Attention and Pixel Attention, respectively.

We use the same method which is used to train CP-Net to train the C-Net and P-Net. The final experimental results are shown in Fig. 6 and Table 2.

Table 2. Quantitative comparisons of Ablation Analysis.

Methods	Indoor		Outdoor	
	PSNR	SSIM	PSNR	SSIM
C-Net	29.98	0.9748	29.00	0.9713
P-Net	28.53	0.9621	28.45	0.9589
CP-Net	31.72	0.9858	30.47	0.9815

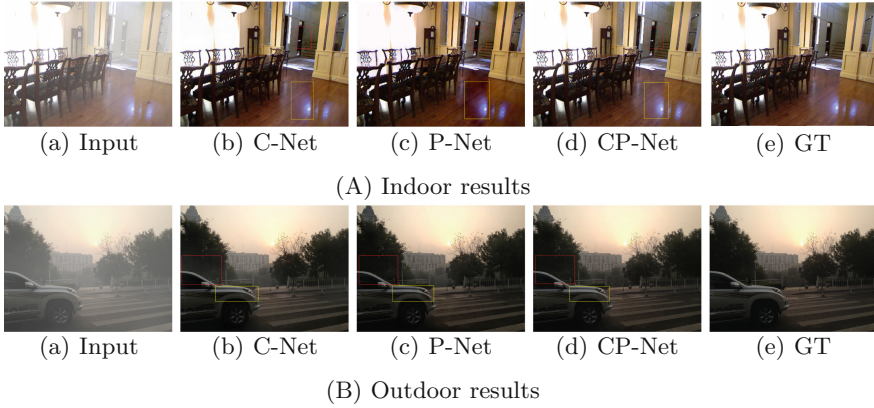


Fig. 6. Qualitative comparisons of Ablation Analysis (color figure online)

Firstly, in the indoor results of Fig. 6, the area marked by the red box contains a large number of black spots in the image recovered by C-Net, but P-Net can restore the area well. In the area marked by the yellow box, P-Net produces a large area of shadow with navy blue light and C-Net can maintain the exact details of the area well. CP-Net can accurately restore these two areas at the same time and generate a de-hazed image closer to the ground truth image.

Secondly, in the outdoor results of Fig. 6, the area marked by the red box contains a lot of haze in the image recovered by C-Net, but P-Net can remove the haze from the area cleanly. In the area marked by the yellow box, P-Net distorts the color, producing a thick black line but C-Net can restore this area well. CP-Net can accurately restore these two areas to produce a clearer image.

In order to further verify the superiority of CP-Net again, we compare CP-Net, C-Net and P-Net with respect to PSNR and SSIM. The comparative results are shown in Table 2. The results indicate that CP-Net is superior to C-Net and P-Net, and has achieved the highest PSNR and SSIM.

The above shows that Channel Attention and Pixel Attention are similar to a complementary relationship, and the fusion of the information captured by them can further enhance the effect of image de-hazing.

6 Conclusion

We propose a new end-to-end single image de-hazing network (CP-Net). Although our network structure is simple, it can surpass many previous state-of-the-art de-hazing methods. CP-Net mainly leverages the combination of channel attention and pixel attention mechanisms. The combination of channel attention and pixel attention treats the information in the feature maps differently, filters out important information, and achieves excellent de-hazing effect.

References

1. Cai, B., Xu, X., Jia, K., et al.: DehazeNet: an end-to-end system for single image haze removal. *IEEE Trans. Image Process.* **25**(11), 5187–5198 (2016)
2. Li, B., Peng, X., Wang, Z., et al.: AOD-net: all-in-one dehazing network. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 4770–4778 (2017)
3. Hide, R.: Optics of the atmosphere: scattering by molecules and particles. *Phys. Bull.* **28**, 521 (1977)
4. Narasimhan, S.G., Nayar, S.K.: Chromatic framework for vision in bad weather. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE (2000)
5. Narasimhan, S.G., Nayar, S.K.: Vision and the atmosphere. *Int. J. Comput. Vis.* **48**(3), 233–254 (2002)
6. He, K., Sun, J., Tang, X., et al.: Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(12), 2341–2353 (2011)
7. Zhu, Q., Mai, J., Shao, L.: A fast single image haze removal algorithm using color attenuation prior. *IEEE Trans. Image Process.* **24**(11), 3522–3533 (2015)
8. Ju, M., Gu, Z., Zhang, D.: Single image haze removal based on the improved atmospheric scattering model. *Neurocomputing* **260**, 180–191 (2017)
9. Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., Yang, M.-H.: Single image dehazing via multi-scale convolutional neural networks. In: *Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9906*, pp. 154–169. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46475-6_10
10. He, K., Zhang, X., Ren, S., et al.: Deep residual learning for image recognition (2015)
11. Li, B., Ren, W., et al.: Benchmarking single-image dehazing and beyond. *IEEE Trans. Image Process.* **28**, 492–505 (2018)
12. Zhang, Y., Li, K., Li, K., et al.: Image super-resolution using very deep residual channel attention networks (2018)
13. Ren, W., Ma, L., Zhang, J., et al.: Gated fusion network for single image dehazing. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3253–3261 (2018)
14. Xu, K., Ba, J., Kiros, R., et al.: Show, attend and tell: neural image caption generation with visual attention. *Computer Science* (2015)
15. Vaswani, A., Shazeer, N., Parmar, N., et al.: Attention is all you need. In: *Advances in Neural Information Processing Systems*, pp. 5998–6008 (2017)
16. Wang, X., Girshick, R., Gupta, A., et al.: Non-local neural networks (2017)
17. He, T., Zhang, Z., Zhang, H., et al.: Bag of tricks for image classification with convolutional neural networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 558–567 (2019)
18. Meng, G., Wang, Y., Duan, J., et al.: Efficient image dehazing with boundary constraint and contextual regularization (2013)
19. Berman, D., Avidan, S.: Non-local image dehazing. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1674–1682 (2016)
20. Guo, T., Li, X., Cherukuri, V., et al.: Dense scene information estimation network for dehazing. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (2019)
21. Li, B., Peng, X., Wang, Z., et al.: End-to-end united video dehazing and detection (2017)

22. Liu, Y., Zhao, G., Gong, B., et al.: Improved techniques for learning to dehaze and beyond: a collective study (2018)
23. Tu, Z., Chen, X., Yuille, A.L., et al.: Image parsing: unifying segmentation, detection, and recognition. *Int. J. Comput. Vis.* **63**(2), 113–140 (2005)
24. Tarel, J.-P., Hautière, N., Cord, A., et al.: Improved visibility of road scene images under heterogeneous fog. In: *IEEE Intelligent Vehicles Symposium*. IEEE (2010)
25. Sakaridis, C., Dai, D., Van Gool, L.: Semantic foggy scene understanding with synthetic data. *Int. J. Comput. Vis.* **126**(9), 973–992 (2018)
26. Fattal, R.: Dehazing using color-lines. *ACM Trans. Graph.* **34**, 1–14 (2014)
27. Jiang, Y., Sun, C., Zhao, Y., et al.: Image dehazing using adaptive bi-channel priors on superpixels. *Comput. Vis. Image Underst.* **165**, 17–32 (2017)
28. Fattal, R.: Single image dehazing. *ACM Trans. Graph.* **27**(3), 1–9 (2008)
29. Russakovsky, O., Deng, J., Su, H., et al.: Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* **115**(3), 211–252 (2015)
30. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. *Computer Science* (2014)
31. Qu, Y., Chen, Y., Huang, J., et al.: Enhanced pix2pix dehazing network. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8160–8168 (2019)
32. Lim, B., Son, S., Kim, H., et al.: Enhanced deep residual networks for single image super-resolution. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 136–144 (2017)