Image and Vision Computing 139 (2023) 104820

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

# Frequency and content dual stream network for image dehazing

Meihua Wang [a], Lei Liao [a], De Huang [a], Zhun Fan [b,*], Jiafan Zhuang [b], Wensheng Zhang [c]

[a] South China Agricultural University, Guangzhou 510642, China
[b] Shantou University, Shantou 515063, China
[c] University of Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

*Keywords:*
Image dehazing
Frequency features
Attention octave convolution
Dual self-attention

ABSTRACT

Image dehazing can improve image clarity and visual effect, which plays a pivotal role in many computer vision tasks. Existing dehazing methods are mostly based on a single feature stream and tend to ignore the low-frequency characteristics of haze. In this paper, we propose a dual stream network for image dehazing. To enhance the edge information and texture detail of the image, we construct a frequency stream based on attention octave convolution. We decompose the features into high and low-frequency branches in the frequency stream to obtain different structural information. By adding a residual channel attention block, the attention octave convolution can extract frequency features more efficiently and effectively. Due to the lower resolution of low-frequency features in the frequency stream, the frequency stream features alone are insufficient for recovering the overall content of the image. Therefore, a content stream was added to compensate for the information lost in the frequency stream. By fusing the outputs of two feature streams, the network achieves an enhanced dehazing performance. The results show that our method is superior to other state-of-the-art algorithms in quantitative evaluation and visual impact.

## 1. Introduction

Adverse weather conditions such as haze and dust can affect image quality, causing loss of contrast and color distortion. Advanced vision tasks, such as object detection [1,2] and image segmentation [3], are prone to degrade significantly when the input image has severe haze. Therefore, dehazing technology is critical in image processing and machine vision.

Physical model-based approaches [4–7] try to remove haze with the help of intermediate variables in the physical models. For example, the classical atmospheric scattering model [8] can be used to recover clean images, including key parameters such as the transmission map and atmospheric light intensity. However, the physical model and prior information can not always reflect the inherent properties of hazy images.

Due to the success of deep learning in various tasks, early non-end-to-end dehazing approaches [9–11] use deep convolutional neural networks to estimate the transmission map and atmospheric light intensity, then dehaze according to the atmospheric scattering model. However, it is challenging to obtain ground truth data of transmission maps. On the other hand, the quality of the dehazed image heavily relies on estimating the intermediate variables.

In recent years, the end-to-end dehazing approaches [12–17] have achieved improved performance. Instead of estimating transmission maps and atmospheric light intensities, these approaches recover clear image directly through powerful feature representation and mapping capabilities of deep convolutional neural networks. Nevertheless, these approaches mainly adopt generic network structures (e.g., GAN [18,19], DenseNet [20], GridNet [21], encoder-decoder networks [22,23]), which limit their dehazing performance due to the fact that they usually extract features in the spatial domain, without taking advantage of features in the frequency domain.

The frequency features of the image contain comprehensive information. High-frequency features correspond to sharp edges and important details of objects, while low-frequency features correspond to information such as overall content and color [24]. Recently, some dehazing methods have also used frequency information to restore haze-free images. Liu et al. [25] used Wavelet Transform to decompose the hazy image into high and low-frequency components, and processed the high and low-frequency components separately to obtain a haze-free image. Xu et al. [26] used the Laplace Operator to obtain high-frequency information of images to improve the quality of dehazed images. These methods requires physical processes to obtain the

* Corresponding author.
*E-mail address:* zfan@stu.edu.cn (Z. Fan).

https://doi.org/10.1016/j.imavis.2023.104820
Received 12 June 2023; Received in revised form 6 September 2023; Accepted 13 September 2023
Available online 16 September 2023

frequency information of the image in advance. In this process, the size of the high-pass and low-pass filters is a crucial parameter that determines the quality of the extracted features. However, for images of different sizes, the filter sizes need to be adjusted accordingly. As a result, using traditional physics-based methods to extract high-frequency and low-frequency features can be highly inefficient in practical applications.

Yu et al. [27] constructed a dual-guided dehazing network based on both frequency and spatial guidance (FSDGN). They obtained the amplitude spectrum and phase spectrum of images through Fourier transformation. After comparing, they found that hazy images and clear images have significant differences in the amplitude spectrum, while the phase spectrum shows minor differences.

Compared with FSDGN, we further analyze the amplitude spectrum of both hazy and clear images. Specifically, we performed a Fourier transform on the hazy image and filtered it using a high-pass filter and a low-pass filter, respectively. The results are shown in Fig. 1. The low-pass filter allows low-frequency information to pass through. After the low-pass filter, the hazy image still has a noticeable haze residue. The results indicate that haze is closer to low-frequency information. The physical distinction between hazy and clear image pairs in the frequency domain is more pronounced than in the spatial domain. Therefore, exploring the correlation between haze degradation and frequency becomes crucial for understanding the dehazing problem. To address the differences between high-frequency and low-frequency images in prior information, we decompose the spatial domain features into high-frequency and low-frequency features in the frequency domain. Our method could extract the high and low frequency features accurately and repair the image with the help of high and low frequency features.

Based on the above analysis, this paper proposes a frequency and content dual stream network for image dehazing. The network learns richer features by building a dual stream network with content and frequency streams. We design a content stream based on a nested residual structure to preserve the overall content of the dehazed image. The frequency stream is decomposed into high and low-frequency branches to provide different structure information. We add residual channel attention to the original octave convolution called attention octave convolution to extract frequency features more accurately. In the skip connections of the frequency stream, we design a dual self-attention (DSA) mechanism to enhance feature communication between high and low-frequency branches. The results demonstrate that our method outperforms other state-of-the-art dehazing algorithms.

The contributions of the paper can be summarized as follows:

1) We propose a frequency and content dual stream network for image dehazing, which learns richer features than single feature stream and restores hazy images from different perspectives.
2) We design a frequency stream to extract the frequency features of hazy images, which we further use attention octave convolution to decompose features into high and low-frequency branches. DSA is proposed to enhance feature communication between high and low-frequency branches. The proposed method uses structural information provided by frequency features to recover details.
3) To compensate for the information lost in the frequency stream, we design a content stream to preserve the overall content of the image. In the content stream, we use the residual channel attention to adaptively adjust the weight of each channel and combine the nested residual structure to filter out the redundant low-frequency information.

## 2. Related works

### 2.1. Image dehazing

Image dehazing methods can be divided into physical model-based and deep learning-based methods. The deep learning-based methods can be divided into non-end-to-end methods and end-to-end methods.

Physical model-based methods use prior information to estimate critical parameters in the model. He et al. [4] proposed the dark channel prior algorithm to get the transmission map through the dark channel map, which achieved a pronounced dehazing effect. Zhu et al. [28] proposed the color attenuation prior algorithm to restore the depth map of the image and then estimate the transmission map. Berman et al. [5] proposed the non-local prior algorithm to estimate the transmission map through the haze-lines. Since prior information is not universally applicable, the application scenarios of these algorithms are limited.

With the rise of deep learning, early non-end-to-end dehazing methods used deep convolutional neural networks to estimate the

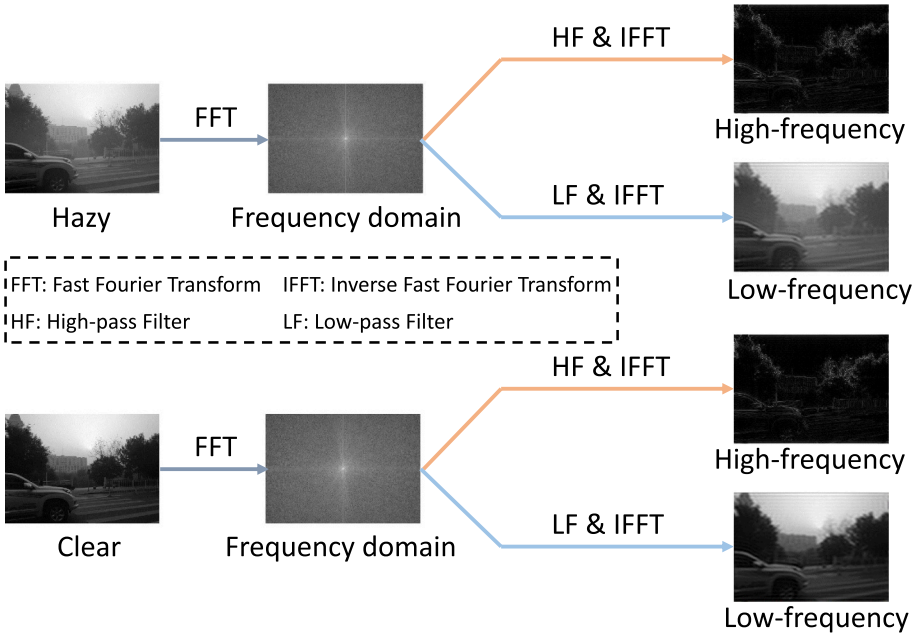


**Fig. 1.** Hazy/clear image frequency decomposed results. No significant difference exists between the hazy and clear images after the high-pass filter. After the low pass filter, the difference is noticeable.

intermediate variables. Cai et al. [29] proposed DehazeNet, which built a network based on prior information in traditional dehazing methods to estimate the transmission map. Ren et al. [9] proposed a multi-scale strategy. They first used a coarse-scale network to estimate the transmission map and then used a fine-scale network to refine the estimation results. Li et al. [10] reconstructed the atmospheric scattering model, integrated the transmission map and atmospheric light intensity into a parameter K, and proposed AODNet to estimate this parameter. Zhang et al. [11] proposed a densely connected pyramid network (DCPDN), which used two sub-networks to estimate transmission map and atmospheric light intensity, respectively, and adopted the discriminator loss of Generative Adversarial Network [30]. However, these non-end-to-end methods are constrained by a simplified physical model. Inaccurate estimation of the transmission map or atmospheric light intensity can have a significant impact on the dehazing results.

Unlike non-end-to-end dehazing methods, end-to-end dehazing methods can directly restore haze-free scenes. Ren et al. [12] proposed a gated fusion network (GFN) based on image fusion, which used a weight map to obtain a weighted fusion of the output images corresponding to the three input images. Zhang et al. [20] proposed a perceptual pyramid deep dehazing network based on dense blocks and residual blocks, and adopted perceptual loss [31] to learn network weights. Liu et al. [21] proposed an attention-based multi-scale network (GDN). The backbone module of GDN is based on attention mechanism, which can effectively exchange information of different scales. Zhao et al. [32] proposed a pyramid global context network (PGC-DN), which learns point-wise long-range dependencies and patch-wise long-range dependencies of hazy images. Dong et al. [33] proposed a multi-scale boosted network with dense feature fusion (MSBDN), which can correct the missing spatial information in high-resolution features. Chen et al. [34] proposed a principled synthetic-to-real dehazing guided by physical priors (PSD), which is fine-tuned in an unsupervised way by using a dehazing network pre-trained on synthetic data as the backbone network. Guo et al. [35] proposed a transformer model (Dehamer) with transmission-aware 3D position embedding and introduced prior information related to haze density. The end-to-end methods have made significant progress in dealing with dehazing problems. However, these methods tend to focus on the global content or local information of hazy images, without fully utilizing the low-frequency characteristics of the haze itself to aid in image restoration. The frequency features of hazy images contains rich information, which helps improve the quality of dehazed image and deserves to be investigated.

### 2.2. Octave convolution

The image can be decomposed into high and low-frequency components, and the feature maps of convolution layer also have high and low-frequency features. The high-frequency features correspond to the area where the intensity values change rapidly, such as the boundaries, edges, and other detailed information. The low-frequency features refer to the area where the intensity values change smoothly, such as the background with the same color and almost the same intensity. In a recent study [36], octave convolution (OctConv) was proposed to process high and low-frequency features separately. OctConv uses a multi-frequency feature representation method that stores and processes low-frequency features by mapping them to low-resolution tensors to reduce redundancy. Unlike the traditional method of separating different frequencies, the high and low-frequency feature maps refer to feature maps with different resolutions. With the intra-frequency update and inter-frequency communication, OctConv separates two kinds of features into two groups of feature maps.

OctConv can also improve the performance of many computer vision tasks by replacing traditional convolution [37–40]. On the image segmentation task, Fan et al. [41] built an accurate retinal vessel segmentation neural network using OctConv and achieved comparable performance to other state-of-the-art methods with a faster processing speed. On the image classification task, Xu et al. [42] proposed a multi-scale octave 3D CNN for hyperspectral image classification, which outperformed many state-of-the-art methods. Up to now, OctConv is rarely used for image dehazing. The purpose of image dehazing is to restore a clean scene and preserve the overall content and textural details. Therefore, we construct a dual stream network, combining a content stream based on traditional convolution and a frequency stream based on OctConv. Our method achieves evaluation results comparable to the latest algorithms, effectively dehazing by leveraging frequency features while preserving the content and detail information of the images.

## 3. Method

The overall structure of the network is shown in Fig. 2, which contains a content stream and a frequency stream. After a $3 \times 3$ convolutional layer, the features will serve as the input for two feature streams. The outputs of the two streams are fused to obtain the final dehazed image.

### 3.1. Content stream based on nested residual structure

Fig. 1 shows that haze is closer to low-frequency information. There is redundant low-frequency information in the hazy image. In addition, frequency stream features will lose a lot of information due to continuous downsampling operations, especially low-frequency features. The frequency stream features alone are insufficient for recovering the overall content of the image. Therefore, we design a content stream based on nested residual structure.

The nested residual structure contains a number of residual groups with long residual connections, with each residual group consisting of a number of residual channel attention blocks with short residual connections. Residual connections allows rich information to be propagated backwards directly through constant mapping, which helps maintain the overall content of the dehazed image. When information is propagated backwards, the network should have the feature discrimination ability to filter out redundant low-frequency information. Therefore, we add the residual channel attention blocks in the nested residual structure. In residual channel attention block, global average pooling and maximal pooling are used to capture global common and distinctive information, respectively. Compared to SENet [43], we use depthwise separable convolution to predict the weight of each channel independently, allowing the channel to use the weight directly and avoiding the dimensionality reduction caused by the fully connected layer. The structure of the residual channel attention block is shown in Fig. 3. It can be described as:

$$F_c = RCA(F) = F + M_c^{*}(W_2(W_1F)) \tag{1}$$

where $F$ and $F_c$ represent input and output features, respectively. $M_c \in R^{C\times1\times1}$ is the channel attention map. $W_1$ and $W_2$ represent the convolution weights of the first two layers, respectively.

$$M_c = \sigma\left(W_p\left(W_dF_{ap}\right) + W_p\left(W_dF_{mp}\right)\right) \tag{2}$$

where $\sigma$ denotes the sigmoid function. $W_d$ and $W_p$ represent the weights of depthwise convolution and pointwise convolution, respectively. $F_{ap}$ and $F_{mp}$ denote the average-pooled features and maximal-pooled features, respectively.

### 3.2. Feature extraction based on attention Octconv

Through Fourier spectrum analysis, we observed significant differences between hazy images and clear images after low-pass filtering, while the differences became minimal after high-pass filtering. In the frequency stream, we decompose the spatial domain features into high-frequency and low-frequency features in the frequency domain.

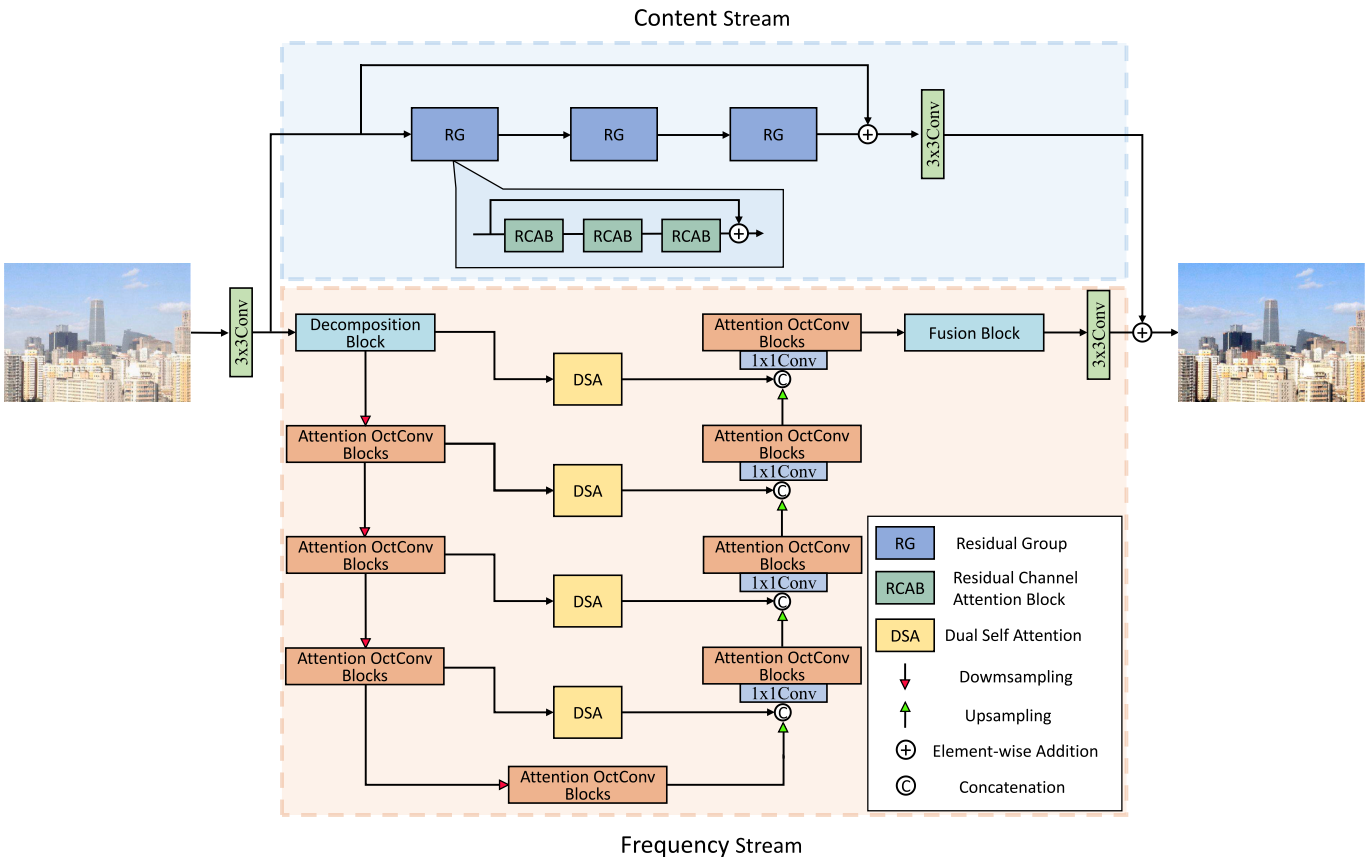


**Fig. 2.** The structure of frequency and content dual stream network.

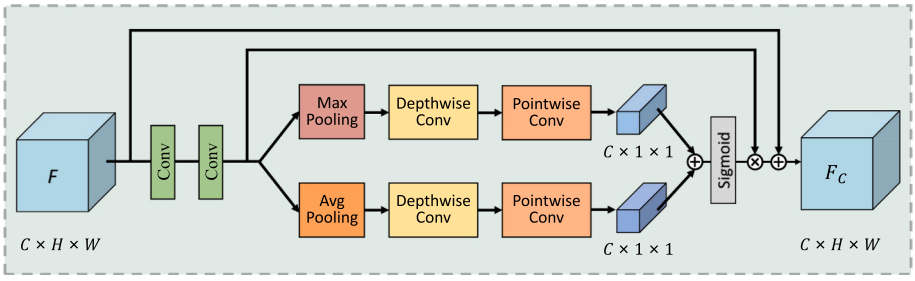


**Fig. 3.** The structure of Residual Channel Attention Block (RCAB).

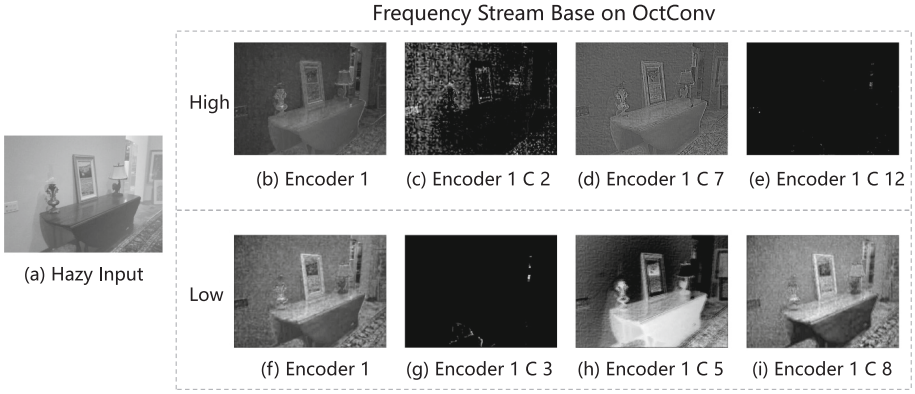


**Fig. 4.** Frequency stream feature visualization based on OctConv. (b) and (f) represent the visualization results of the high-frequency and low-frequency features, respectively. (c)-(e) and (g)-(i) respectively show the visualization results of single channel feature map. For instance, Encoder 1 means the output features of the first encoder, and Encoder 1 C 2 means the features of the second channel in the output features of the first encoder.

OctConv [36] divides features into high and low-frequency branches by channels. It uses feature maps with different resolutions to represent high and low-frequency features. Particularly, OctConv is able to adjust the ratio of high and low-frequency components. For the redundant low-frequency information, we set the ratio of low-frequency components in the frequency stream to 0.25 and the ratio of high-frequency components to 0.75. As the low-frequency information contains more haze characteristics, reducing the proportion of low-frequency information in the frequency stream can directly and effectively dehaze the image. At the same time, the network will focus more on high-frequency information, leading to a restored image that closely resembles the clear image.

However, OctConv is not entirely accurate in extracting frequency features. As shown in Fig. 4, we performed feature visualization on the OctConv-based frequency stream. Taking the high-frequency branch as an example, (b) represents the overall feature visualization result of all channels, (c)-(e) represent the feature visualization results of different channels. We can see that different channels contain different information. For instance, high-frequency information (d), noises (c), and irrelevant information (e). These noises and irrelevant information will affect the extraction of frequency features in the frequency stream.

To reduce the interference of noises and irrelevant information, we propose an Attention OctConv based on residual channel attention, as shown in Fig. 5. The Attention OctConv adds residual channel attention before performing feature exchange and update between high and low-frequency branches. By adjusting each channel's weight, the network will focus more on frequency features, thus achieving more accurate feature extraction. The Attention OctConv can be expressed as:

$$\begin{cases} Y^H = Y^{H\to H} + Y^{L\to H} \\ Y^L = Y^{L\to L} + Y^{H\to L} \end{cases} \tag{3}$$

where $Y^H$ and $Y^L$ represent the high and low-frequency branches of the output, respectively. $Y^{H\to H}$ indicates the mapping between the high-frequency and high-frequency branches. $Y^{L\to H}$ indicates the mapping between the low-frequency and high-frequency branches. $Y^{L\to L}$ indicates the mapping between the low-frequency and low-frequency branches. $Y^{H\to L}$ indicates the mapping between the high-frequency and low-frequency branches.

$$\begin{cases} Y^{H\to H} = f\left(RCA\left(X^H\right); W^{H\to H}\right) \\ Y^{L\to H} = Upsampling\left(f\left(RCA\left(X^L\right); W^{L\to H}\right), 2\right) \end{cases} \tag{4}$$

where $X^H$ and $X^L$ represent the high and low-frequency branches of the input, respectively. $RCA(\cdot)$ represents residual channel attention operation. $f(X; W)$ represents the convolution operation with input $X$ and convolution kernel $W$. $W^{H\to H}$ denotes the convolution kernel from high-frequency branch to the high-frequency branch. $W^{L\to H}$ denotes the convolution kernel from low-frequency branch to the high-frequency branch. $Upsampling(X, 2)$ denotes the upsampling operation with input $X$ and the sampling factor is 2.

$$\begin{cases} Y^{L\to L} = f\left(RCA\left(X^L\right); W^{L\to L}\right) \\ Y^{H\to L} = f\left(Pooling\left(RCA\left(X^H\right), 2\right); W^{H\to L}\right) \end{cases} \tag{5}$$

where $W^{L\to L}$ denotes the convolution kernel from low-frequency branch to low-frequency branch. $W^{H\to L}$ denotes the convolution kernel from high-frequency branch to low-frequency branch. $Pooling(X, 2)$ denotes the pooling operation with input $X$ and the stride is 2.

To decompose the features extracted by traditional convolution in the initial stage and fuse the features extracted by Attention OctConv in the final stage, we design a decomposition block and a fusion block, as shown in Fig. 6. Unlike CBAM [44], we use depthwise separable convolution to generate the spatial attention map that preserves location information and avoids information loss caused by the pooling layer. The decomposition block preliminarily divides features into high and low-frequency branches, which helps the frequency stream to extract frequency features. The fusion block fuses the frequency features of the high and low-frequency branches, which helps to enhance the details of the dehazed image.

### 3.3. Dual self-attention

The frequency stream is designed with a U-Net structure, which may result in a loss of substantial spatial information during the encoding stage. To supplement the information and enhance image recovery during the decoding phase, we have incorporated the dual self-attention (DSA) module into the skip connections of the same-level features. DSA module enhances the high and low-frequency features from the encoding stage and plays a guiding and complementary role during the decoding stage, as shown in Fig. 7.

We upsample the features of the low-frequency branch to the same resolution as the high-frequency branch and then perform unified processing to enhance the feature communication. Inspired by self-attention in transformer [49,50], we apply the self-attention mechanism to the frequency features.

$$\begin{cases} Q = W_2^Q W_1^Q X \\ K = W_2^K W_1^K X \\ V = W_2^V W_1^V X \end{cases} \tag{6}$$

where $X$ represents the frequency features after concatenation. $W_1^{(\cdot)}$ and $W_2^{(\cdot)}$ denote $1 \times 1$ convolution and $3 \times 3$ convolution, respectively. $Q$, $K$, and $V$ represent the projections of Query, Key, and Value, respectively.

Finally, DSA generates spatial attention maps with two channels. The spatial attention maps are split by channel and multiplied with the features of high and low-frequency branches separately.

$$\widehat{X} = \sigma W_1(\text{Attention}(Q, K, V) + X) \tag{7}$$

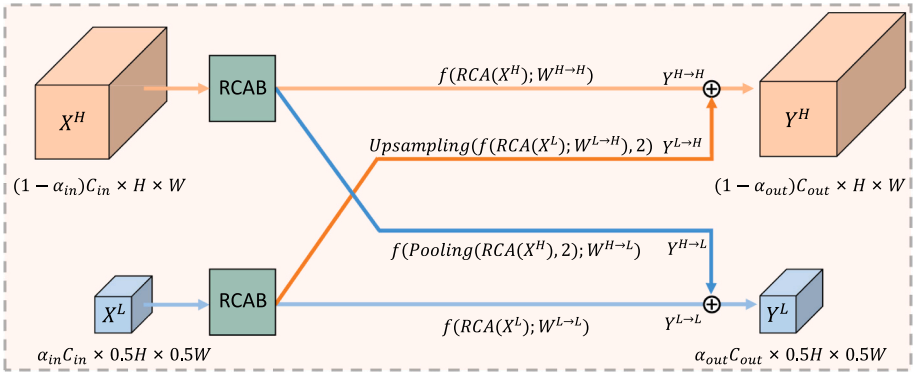


**Fig. 5.** The structure of Attention OctConv. Here, we set $\alpha_{in} = \alpha_{out} = 0.25$.

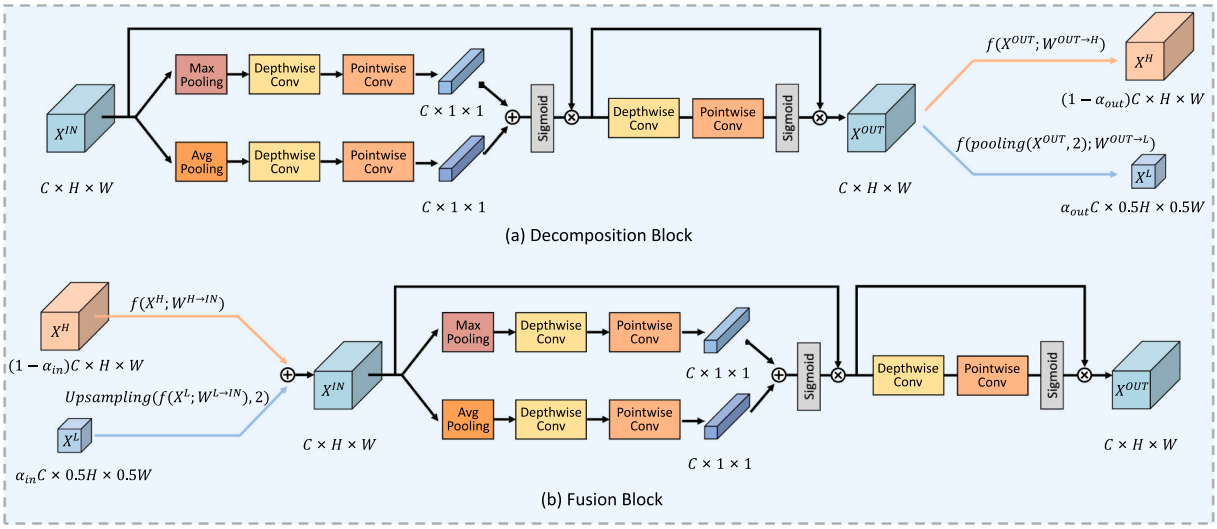


**Fig. 6.** The structure of decomposition block and fusion block. (a) is the decomposition block and (b) is the fusion block.

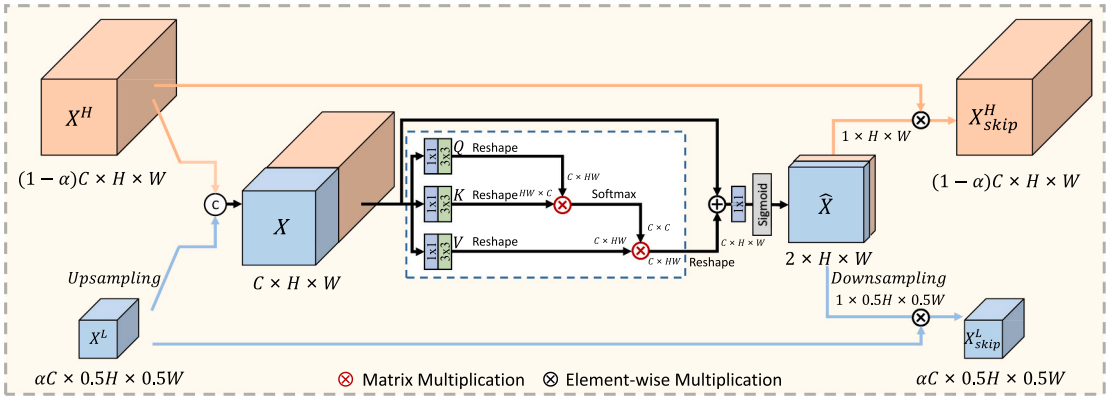


**Fig. 7.** The structure of dual self-attention block. The white dashed box shows the self-attention mechanism.

$$\text{Attention}(Q,K,V) = \text{Softmax}\left(QK^T\right)V \tag{8}$$

where $\widehat{X}$ represents the spatial attention maps. $\sigma$ denotes the sigmoid function.

In the process described above, DSA captures the internal associations between high and low frequency features, and plays a role in feature enhancement. The high and low-frequency features generated by each encoder layer are then fused with the corresponding decoder features after being enhanced by DSA.

## 4. Experiment

To verify the superiority of the proposed method, we compared our method with other methods on both synthetic and real-world hazy datasets, as well as some locally obtained images. Then, we conducted an ablation analysis to demonstrate the effectiveness of the core modules used in the proposed method.

### 4.1. Training settings

The method was based on the PyTorch framework, and all experiments were performed on a single NVIDIA GeForce RTX 3090 GPU. We use ADAM with $\beta_1 = 0.9$, $\beta_2 = 0.999$ for optimization. The initial learning rate is 0.0001, and the learning rate is adjusted by the cosine annealing strategy. The batch and total number of iterations are 16 and 400 k, respectively. We use synthetic indoor dataset ITS and outdoor dataset OTS as the training set and SOTS [51] as the testing set. We also train and test on real-world datasets I-HAZE [52], O-HAZE [53], and DENSE-HAZE [54]. Furthermore, we test some real-world hazy images using the model trained on the OTS dataset. In the training phase, 256 × 256 patches are cropped randomly from the hazy images and randomly flipped horizontally after normalization.

We compare our method with the SOTA methods, including DCP [4] (TPIMA'2010), AOD-Net [10] (ICCV'2017), GDN [21] (ICCV'2019), PGC-DN [32] (TCSVT'2020), FFA-Net [15] (AAAI'2020), MSBDN [33] (CVPR'2020), AECR-Net [45] (CVPR'2021), Dehamer [35] (CVPR'2022), MAXIM-2S [46] (CVPR'2022), FSDGN [27] (ECCV'2022), CARL-Net [47] (IJCAI'2022), TUSR-Net [48] (TIP'2023). We employ commonly-used PSNR (dB) and SSIM to quantify the dehazing performance of different methods.

### 4.2. Evaluation

Table 1 shows the quantitative results of all the methods in the above five datasets. We have obtained 36.39 dB PSNR and 0.9871 SSIM on SOTS (Indoor) dataset, 34.45 dB PSNR and 0.9851 SSIM on SOTS (Outdoor) dataset. Although the performance of our method is not outstanding on the SOTS (Indoor) dataset, it excels on the SOTS (Outdoor) dataset, achieving the second-best results. Our proposed method achieves 26.76 dB PSNR and 0.8670 SSIM on I-HAZE, 24.19 dB PSNR and 0.8639 SSIM on O-HAZE, 16.85 dB PSNR and 0.5201 SSIM on DENSE-HAZE. In particular, our method achieves much higher SSIM

**Table 1**
Comparison of performance on public datasets. Red texts and blue texts indicate the best and the second-best performance respectively.

| Method | SOTS (Indoor) | | SOTS (Outdoor) | | I-HAZE | | O-HAZE | | DENSE-HAZE | |
|---|---|---|---|---|---|---|---|---|---|---|
| | PSNR ↑ | SSIM ↑ | PSNR ↑ | SSIM ↑ | PSNR ↑ | SSIM ↑ | PSNR ↑ | SSIM ↑ | PSNR ↑ | SSIM ↑ |
| DCP [4] (TPIMA'2010) | 16.62 | 0.8179 | 19.13 | 0.8148 | 14.43 | 0.7520 | 16.78 | 0.6530 | 10.06 | 0.3856 |
| AOD-Net [10] (ICCV'2017) | 20.51 | 0.8162 | 24.14 | 0.9198 | 13.98 | 0.7320 | 17.56 | 0.6500 | 11.57 | 0.3959 |
| GDN [21] (ICCV'2019) | 32.16 | 0.9836 | 30.86 | 0.9819 | 16.62 | 0.7870 | 18.92 | 0.6720 | 13.31 | 0.3681 |
| PGC-DN [32] (TCSVT'2020) | – | – | 28.61 | 0.9510 | 26.99 | 0.8890 | 24.91 | 0.7730 | – | – |
| FFA-Net [15] (AAAI'2020) | 36.39 | 0.9886 | 33.57 | 0.9840 | – | – | – | – | 14.39 | 0.4524 |
| MSBDN [33] (CVPR'2020) | 32.77 | 0.9813 | 34.29 | 0.9850 | 23.93 | 0.8910 | 24.36 | 0.7490 | 15.37 | 0.4858 |
| AECR-Net [45] (CVPR'2021) | 37.17 | 0.9901 | – | – | – | – | – | – | 15.80 | 0.4660 |
| Dehamer [35] (CVPR'2022) | 36.63 | 0.9881 | 35.18 | 0.9860 | – | – | – | – | 16.62 | 0.5602 |
| MAXIM-2S [46] (CVPR'2022) | 38.11 | 0.9910 | 34.19 | 0.9850 | – | – | – | – | – | – |
| FSDGN [27] (ECCV'2022) | 38.63 | 0.9903 | – | – | – | – | – | – | 16.91 | 0.5806 |
| CARL-Net [47] (IJCAI'2022) | 41.92 | 0.9954 | 33.26 | 0.9849 | 25.43 | 0.8807 | 25.83 | 0.8078 | 15.47 | 0.5482 |
| TUSR-Net [48] (TIP'2023) | 38.67 | 0.9911 | – | – | – | – | 25.34 | 0.7656 | 18.62 | 0.5606 |
| Ours | 36.39 | 0.9871 | 34.45 | 0.9851 | 26.76 | 0.8670 | 24.19 | 0.8639 | 16.85 | 0.5201 |

value on the real-world dataset O-HAZE than the other algorithms. The results show that our method has comparable dehazing performance to state-of-the-art algorithms.

Fig. 8 shows the dehazed visual comparison of the synthetic datasets SOTS (Indoor) and SOTS (Outdoor). DCP have problem with color distortion in indoor and outdoor dehazed images. AOD-Net does not remove the haze completely, and there is still a lot of haze in the last row of dehazed results. GDN, FFA-Net and MSBDN show comparable dehazing performance to our method on synthetic datasets. The restored images all have normal color and no obvious haze residue.

Fig. 9 shows the dehazed results of the real-world datasets I-HAZE, O-HAZE and DENSE-HAZE. DCP has noticeable dehazing effects in the indoor dehazed image, but the outdoor dehazed image appears to have severe color distortion. The dehazed images of AOD-Net have noticeable haze residue. GDN and MSBDN perform well on synthetic datasets. However, their performance on real datasets is inferior to our method. The dehazed results of GDN have noticeable artifacts. The enlarged details of MSBDN show apparent color deviation compared to our method. Our method performs better in haze removal and color recovery than other methods.

Fig. 10 shows the visual comparison of real-world hazy images. DCP have problem with color distortion in the dehazed images. The enlarged details of GDN, FFA-Net and MSBDN in the first row have noticeable haze residue. In the last row, the dehazed results from AOD-Net, GDN, and MSBDN show an unnatural color for the sky. Compared with other methods, our method has an apparent dehazing effect while preserving the original color.

### 4.3. Ablation experiment

We perform ablation experiments to verify the effectiveness of Attention OctConv, Content Stream and DSA in the model. After removing and replacing the corresponding modules, the ablation models are obtained using the same training strategy. We used ITS as the training set and SOTS (Indoor) as the testing set. Table 2 shows the results of the ablation experiment.

M0 is the base model, using Original OctConv, no Content Stream, and no DSA. M1 uses only Attention OctConv to extract frequency features, no Content Stream, and no DSA. From the results of M0 and M1, we can see that Attention OctConv performs better than the original OctConv. M2 uses original OctConv to extract frequency features and adds DSA for feature enhancement of high and low-frequency branches. Due to the lack of the content stream to supplement the global content information, the performances of M1 and M2 are not so satisfactory. M3 adds the content stream, which forms a dual stream network with the frequency stream. By combining the two feature streams, M3 learns

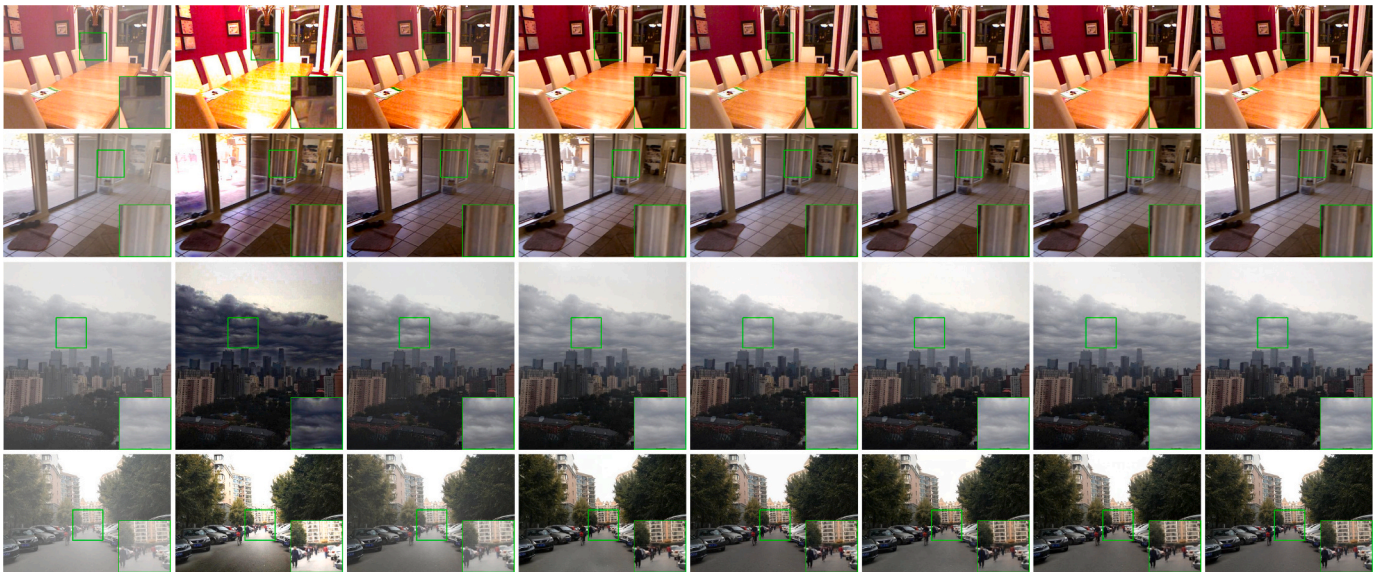


**Fig. 8.** Comparison of dehazing results on SOTS (Indoor) and SOTS (Outdoor) datasets. The first two rows of hazy images are from SOTS (Indoor), and the last two rows are from SOTS (Outdoor).

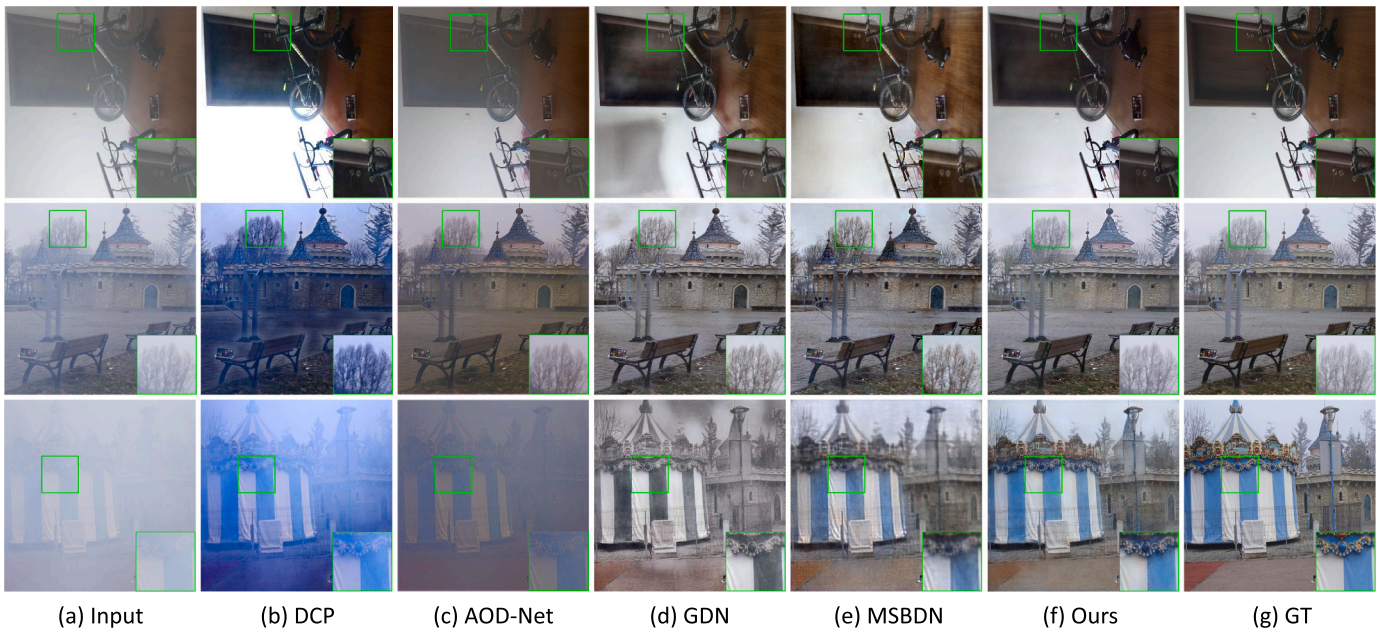


**Fig. 9.** Comparison of dehazing results on I-HAZE, O-HAZE and DENSE-HAZE datasets. The first row is from I-HAZE, the second row is from O-HAZE, and the last row is from DENSE-HAZE.

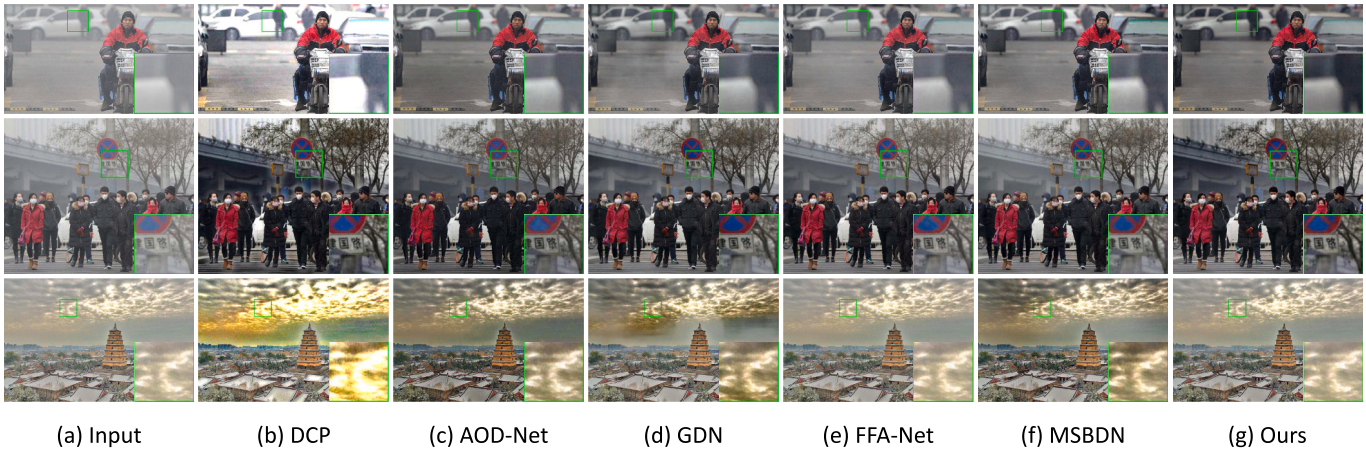


**Fig. 10.** Comparison of dehazing results on real-world hazy images. Note that these hazy images are from the Unannotated Real-World Hazy Images in RESIDE and have no ground truth.

**Table 2**
Ablation experiment results on SOTS (Indoor) dataset. M0-M7 total 8 different conbinations.

| Module | M0 | M1 | M2 | M3 | M4 | M5 | M6 | M7(ours) |
|---|---|---|---|---|---|---|---|---|
| Attention OctConv | ✘ | ✓ | ✘ | ✘ | ✓ | ✓ | ✘ | ✓ |
| DSA | ✘ | ✘ | ✓ | ✘ | ✓ | ✘ | ✓ | ✓ |
| Content Stream | ✘ | ✘ | ✘ | ✓ | ✘ | ✓ | ✓ | ✓ |
| PSNR ↑ | 34.72 | 35.19 | 35.08 | 35.36 | 35.70 | 35.96 | 35.84 | **36.39** |
| SSIM ↑ | 0.9778 | 0.9804 | 0.9792 | 0.9811 | 0.9832 | 0.9850 | 0.9843 | **0.9871** |
| Parameters | 8.95 M | 13.28 M | 10.01 M | 8.98 M | 14.35 M | 13.31 M | 10.04 M | 14.38 M |

richer features and significantly improves performance. Compared with M1, M4 further improves the dehazing performance due to DSA. DSA effectively enhances the feature communication between the high and low-frequency branches. Compared with M1, M5 significantly improves its dehazing performance after integrating supplementary information from the content stream. Due to the lack of Attention OctConv to extract high and low-frequency branches more accurately, the performance of the M6 still has room to be improved. M7 integrates all modules and achieves the best dehazing performance. The results of the ablation experiments show that each module plays an irreplaceable role in the network.

Furthermore, we compared RCAB with the channel attention in the classical SENet, and the results are shown in Table 3. The results indicate a clear advantage of RCAB over the channel attention in SENet. RCAB is used in both the content stream and the frequency stream. On the one hand, the RCAB module reduces the influence of noise and irrelevant

**Table 3**
The comparison of different channel attention module.

| Settings | SOTS (Indoor) | |
|---|---|---|
| | PSNR ↑ | SSIM ↑ |
| Ours + SE Attention | 35.94 | 0.9847 |
| Ours + RCAB | **36.39** | **0.9871** |

information by adjusting the channel weights, which helps the frequency stream to extract frequency features more accurately. On the other hand, the content stream uses RCAB to filter out redundant information, helping to dehaze while preserving content information.

#### 4.3.1. *Effectiveness of the attention OctConv*

We show that Attention OctConv can extract frequency features more accurately than original OctConv. We respectively visualized the frequency stream features of M6 and M7 in the ablation experiment, as shown in Fig. 12. Both M6 and M7 contain the content stream and DSA. The difference is that M6 uses original OctConv, while M7 uses Attention OctConv. As shown in Fig. 12, the high-frequency features of the desk edge and the sofa texture are more precise on the right (after Attention OctCov) than on the left (after original OctCov). In addition, we visualized the features of each stage of the frequency stream in M6 and M7, as shown in Fig. 11. Taking high-frequency features as an example, the results of Attention OctConv are more accurate on the edges and contours of the desk. The feature visualization results show that Attention OctConv can extract frequency features more accurately than the original OctConv.

#### 4.3.2. *Effectiveness of the content stream*

We also show that the content stream and frequency stream focus on different features. Fig. 13 shows the output comparison of the content stream and frequency stream. The output of the content stream removes some of the haze while preserving the overall content of the image. However, the edges and details are unclear. Unlike the content stream, the frequency stream learns more frequency features, and the edges and contours in the image are more clearly identified. Our method maintains the overall content through the content stream and enhances the edges and details through the frequency stream. The final output is closer to the ground truth image by fusing the features of the two streams.

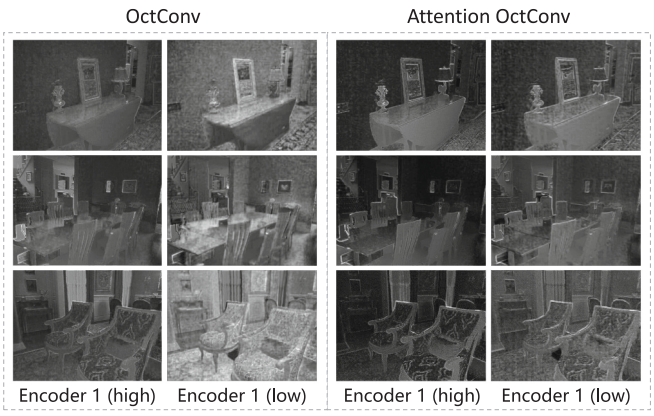


**Fig. 12.** Comparison of feature visualization between Attention OctConv and OctConv.

### 4.4. *Underwater image restoration*

The hazy image is affected by particles suspended in the air. The underwater image restoration problem resembles the image dehazing problem. Due to the scattering and absorption of light by water, the underwater image appears to have color deviation. Therefore, many dehazing algorithms are used for underwater image restoration, such as the Retinex and DCP algorithms. We apply the proposed method to underwater image restoration and compare it with other algorithms.

We select the dataset EUVP [55] widely used in underwater image restoration. The EUVP contains 2185 images. We make the first 2000 images as the training set and the last 185 images as the testing set. The training settings are as previously described. We also use PSNR and SSIM to evaluate the restored images quantitatively. Table 4 shows the quantitative results. Fig. 14 shows the qualitative results.

We have selected several dehazing algorithms for comparison. In

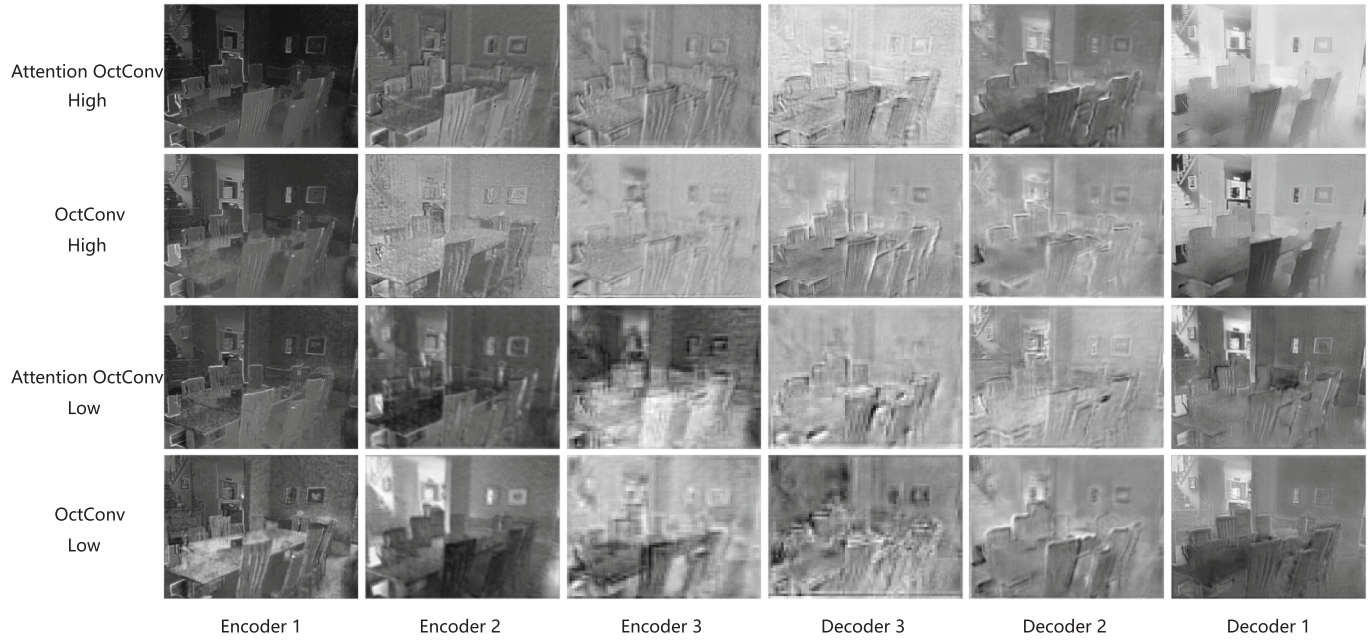


**Fig. 11.** Visual comparison of frequency stream features at different stages. The first two rows are the high-frequency features visualization results extracted by Attention OctConv and OctConv, respectively. The last two rows are the low-frequency features visualization results extracted by Attention OctConv and OctConv, respectively.


**Fig. 13.** Comparison of content stream and frequency stream outputs.

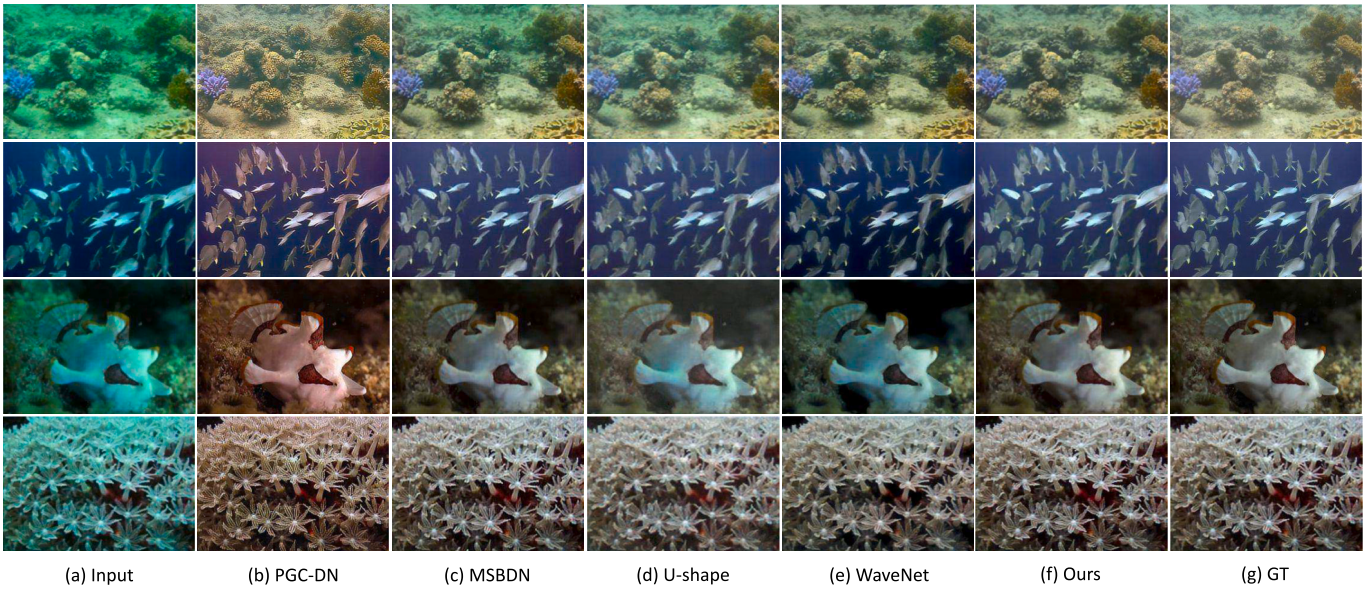


**Fig. 14.** Visual comparison of underwater image enhancement dataset.

**Table 4**
Results on underwater image enhancement dataset.

| Method | Test-E185 | |
|---|---|---|
| | PSNR ↑ | SSIM ↑ |
| PGC-DN [32] (TCSVT'2020) | 22.83 | 0.8304 |
| MSBDN [33] (CVPR'2020) | 27.91 | 0.8666 |
| Dehamer [35] (CVPR'2022) | 28.50 | 0.8906 |
| U-shape [56] (TIP'2023) | 28.75 | 0.8825 |
| WaveNet [57] (TOMM'2023) | 28.62 | 0.8391 |
| Ours | **28.81** | **0.8911** |

addition, U-shape [56] and WaveNet [57] are the latest algorithms in underwater image restoration. Table 4 shows our method performs best on the EUVP dataset. In Fig. 14, the results of PGC-DN and MSBDN show apparent color deviation. The results of U-shape and WaveNet have better color recovery, but the quality and detail of the images are poorer. In the last row, the texture details of the U-shape result are unclear, and the color of the WaveNet result is dark. Compared with other methods, our method achieves the best visual performance. Because the overall color of the underwater images is green or blue, the color deviation is closer to low-frequency information. Our method has specially designed mechanisms of extracting frequency features, which can better repair color deviation. The results show that our method not only produces better dehazed images, but also performs well in underwater image restoration.

## 5. Conclusion

In this paper, we propose a frequency and content dual stream network for single image dehazing. By introducing a dual stream structure, the network can learn richer features and restore images from different perspectives. In the frequency stream, we use attention octave convolution to extract frequency features more accurately. In addition, DSA is designed to enhance the feature communication of high and low-frequency branches. To compensate for the content information lost in the frequency stream, we add the content stream to preserve the overall content of the image. The dual stream network effectively fuses features from two streams to improve the quality of dehazed images. The ablation experiments show that the proposed modules are effective in image dehazing. Comprehensive experiments show that the proposed model outperforms other state-of-the-art methods in image dehazing and underwater image enhancement.

In future work, we will strive to improve the generality of the model

for image restoration in other adversarial weather conditions.

## Credit authorship contribution statement

**Meihua Wang**:Conceptualization, Methodology, Review editing, Supervision. **Lei Liao**:Software, Writing original draft. **De Huang**: Writing original draft. **Zhun Fan**:Writing review, Editing. **Jiafan Zhuang**:Writing review, Editing. **Wensheng Zhang**: Investigation.

## Statement

We have carefully considered the reviewers' comments in revising the manuscript. We sincerely thank the editors and reviewers for their constructive comments, all of which are very valuable and helpful in improving the quality of this paper.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under grants 61976052 and 62176147, and in part by the Key Program of Wen's Joint Fund of Guangdong Basic and Applied Basic Research Fund under grant 2019B1515210009.

## References

[1] B. Li, X. Peng, Z. Wang, J. Xu, D. Feng, End-to-end united video dehazing and detection, in: Proceedings of the AAAI Conference on Artificial Intelligence 32, 2018, pp. 7016–7023.
[2] Y. Chen, W. Li, C. Sakaridis, D. Dai, L. Van Gool, Domain adaptive faster r-cnn for object detection in the wild, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 3339–3348.
[3] C. Sakaridis, D. Dai, L. Van Gool, Semantic foggy scene understanding with synthetic data, Int. J. Comput. Vis. 126 (9) (2018) 973–992.
[4] K. He, J. Sun, X. Tang, Single image haze removal using dark channel prior, IEEE Trans. Pattern Anal. Mach. Intell. 33 (12) (2010) 2341–2353.
[5] D. Berman, S. Avidan, et al., Non-local image dehazing, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1674–1682.
[6] J. Wang, K. Lu, J. Xue, N. He, L. Shao, Single image dehazing based on the physical model and msrcr algorithm, IEEE Trans. Circuits Syst. Video Technol. 28 (9) (2017) 2190–2199.
[7] M. Ju, C. Ding, W. Ren, Y. Yang, Idbp: image dehazing using blended priors including non-local, local, and global priors, IEEE Trans. Circuits Syst. Video Technol. 32 (7) (2021) 4867–4871.
[8] S.G. Narasimhan, S.K. Nayar, Vision and the atmosphere, Int. J. Comput. Vis. 48 (3) (2002) 233–254.
[9] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, M.-H. Yang, Single image dehazing via multi-scale convolutional neural networks, in: European Conference on Computer Vision, Springer, 2016, pp. 154–169.
[10] B. Li, X. Peng, Z. Wang, J. Xu, D. Feng, Aod-net: All-in-one dehazing network, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 4770–4778.
[11] H. Zhang, V.M. Patel, Densely connected pyramid dehazing network, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 3194–3203.
[12] 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 IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 3253–3261.
[13] D. Engin, A. Genç, H. Kemal Ekenel, Cycle-dehaze: Enhanced cyclegan for single image dehazing, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018, pp. 825–833.
[14] 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: 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE, 2019, pp. 1375–1383.
[15] 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 34, 2020, pp. 11908–11915.
[16] X. Zhang, T. Wang, W. Luo, P. Huang, Multi-level fusion and attention-guided cnn for image dehazing, IEEE Trans. Circuits Syst. Video Technol. 31 (11) (2020) 4162–4173.
[17] X. Zhang, J. Wang, T. Wang, R. Jiang, Hierarchical feature fusion with mixed convolution attention for single image dehazing, IEEE Trans. Circuits Syst. Video Technol. 32 (2) (2021) 510–522.
[18] R. Li, J. Pan, Z. Li, J. Tang, Single image dehazing via conditional generative adversarial network, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 8202–8211.
[19] P. Wang, H. Zhu, H. Huang, H. Zhang, N. Wang, Tms-Gan: a twofold multi-scale generative adversarial network for single image dehazing, IEEE Trans. Circuits Syst. Video Technol. 32 (5) (2021) 2760–2772.
[20] H. Zhang, V. Sindagi, V.M. Patel, Multi-scale single image dehazing using perceptual pyramid deep network, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018, pp. 902–911.
[21] X. Liu, Y. Ma, Z. 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.
[22] P. Li, J. Tian, Y. Tang, G. Wang, C. Wu, Deep retinex network for single image dehazing, IEEE Trans. Image Process. 30 (2020) 1100–1115.
[23] D. Zhao, L. Xu, L. Ma, J. Li, Y. Yan, Pyramid global context network for image dehazing, IEEE Trans. Circuits Syst. Video Technol. 31 (8) (2020) 3037–3050.
[24] M. Cai, H. Zhang, H. Huang, Q. Geng, Y. Li, G. Huang, Frequency domain image translation: More photo-realistic, better identity-preserving, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 13930–13940.
[25] F. Liu, L. Cao, X. Shao, P. Han, X. Bin, Polarimetric dehazing utilizing spatial frequency segregation of images, Appl. Opt. 54 (27) (2015) 8116–8122.
[26] Y. Xu, Y. Zhang, Z. Li, Z. Cui, Y. Yang, Multi-scale dehazing network via high-frequency feature fusion, Comput. Graph. 107 (2022) 50–59.
[27] H. Yu, N. Zheng, M. Zhou, J. Huang, Z. Xiao, F. Zhao, Frequency and spatial dual guidance for image dehazing, in: European Conference on Computer Vision, Springer, 2022, pp. 181–198.
[28] Q. Zhu, J. Mai, L. Shao, A fast single image haze removal algorithm using color attenuation prior, IEEE Trans. Image Process. 24 (11) (2015) 3522–3533.
[29] B. Cai, X. Xu, K. Jia, C. Qing, D. Tao, Dehazenet: An end-to-end system for single image haze removal, IEEE Trans. Image Process. 25 (11) (2016) 5187–5198.
[30] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial networks, Commun. ACM 63 (11) (2020) 139–144.
[31] J. Johnson, A. Alahi, L. Fei-Fei, Perceptual losses for real-time style transfer and super-resolution, in: European Conference on Computer Vision, Springer, 2016, pp. 694–711.
[32] D. Zhao, L. Xu, L. Ma, J. Li, Y. Yan, Pyramid global context network for image dehazing, IEEE Trans. Circuits Syst. Video Technol. 31 (8) (2020) 3037–3050.
[33] H. Dong, J. Pan, L. Xiang, Z. Hu, X. Zhang, F. Wang, M.-H. Yang, Multi-scale boosted dehazing network with dense feature fusion, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 2157–2167.
[34] 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.
[35] C.-L. Guo, Q. Yan, S. Anwar, R. Cong, W. Ren, C. 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.
[36] Y. Chen, H. Fan, B. Xu, Z. Yan, Y. Kalantidis, M. Rohrbach, S. Yan, J. Feng, Drop an octave: Reducing spatial redundancy in convolutional neural networks with octave convolution, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 3435–3444.
[37] L. Bai, Q. Liu, C. Li, Z. Ye, M. Hui, X. Jia, Remote sensing image scene classification using multiscale feature fusion covariance network with octave convolution, IEEE Trans. Geosci. Remote Sens. 60 (2022) 1–14.
[38] Z. Kan, S. Li, M. Hou, L. Fang, Y. Zhang, Attention-based octave network for hyperspectral image denoising, IEEE J. Select. Top. Appl. Earth Observ. Remote Sens. 15 (2021) 1089–1102.
[39] C.-M. Feng, Z. Yang, H. Fu, Y. Xu, J. Yang, L. Shao, Donet: dual-octave network for fast mr image reconstruction, IEEE Trans. Neural Netw. Learn. Syst. (2021) 1–11, https://doi.org/10.1109/TNNLS.2021.3090303.
[40] Z. Wang, F. Chen, H. Cheng, Anti-noise object tracking based on siamese octave convolution and attentional correlation fusion, in: 2021 IEEE 6th International Conference on Signal and Image Processing (ICSIP), IEEE, 2021, pp. 670–674.
[41] Z. Fan, J. Mo, B. Qiu, W. Li, G. Zhu, C. Li, J. Hu, Y. Rong, X. Chen, Accurate retinal vessel segmentation via octave convolution neural network, arXiv (2019) preprint arXiv:1906.12193.
[42] Q. Xu, Y. Xiao, D. Wang, Csa-mso3dcnn: multiscale octave 3d cnn with channel and spatial attention for hyperspectral image classification, Remote Sens. 12 (1) (2020) 188–211.
[43] J. Hu, L. Shen, G. Sun, Squeeze-and-excitation networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 7132–7141.
[44] S. Woo, J. Park, J.-Y. Lee, I.S. Kweon, Cbam: Convolutional block attention module, in: Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 3–19.

[45] H. Wu, Y. Qu, S. Lin, J. Zhou, R. Qiao, Z. Zhang, Y. Xie, L. Ma, Contrastive learning for compact single image dehazing, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 10551–10560.
[46] Z. Tu, H. Talebi, H. Zhang, F. Yang, P. Milanfar, A. Bovik, Y. Li, Maxim: Multi-axis mlp for image processing, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 5769–5780.
[47] D. Cheng, Y. Li, D. Zhang, N. Wang, X. Gao, J. Sun, Robust single image dehazing based on consistent and contrast-assisted reconstruction, Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (2022) 848–854.
[48] X. Song, D. Zhou, W. Li, Y. Dai, Z. Shen, L. Zhang, H. Li, Tusr-net: triple unfolding single image dehazing with self-regularization and dual feature to pixel attention, IEEE Trans. Image Process. 32 (2023) 1231–1244.
[49] H. Chen, Y. Wang, T. Guo, C. Xu, Y. Deng, Z. Liu, S. Ma, C. Xu, C. Xu, W. Gao, Pre-trained image processing transformer, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 12299–12310.
[50] S.W. Zamir, A. Arora, S. Khan, M. Hayat, F.S. Khan, M.-H. Yang, Restormer: efficient transformer for high-resolution image restoration, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 5728–5739.
[51] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, Z. Wang, Benchmarking single-image dehazing and beyond, IEEE Trans. Image Process. 28 (1) (2018) 492–505.
[52] C. Ancuti, C.O. Ancuti, R. Timofte, C. De Vleeschouwer, I-haze: a dehazing benchmark with real hazy and haze-free indoor images, in: International Conference on Advanced Concepts for Intelligent Vision Systems, Springer, 2018, pp. 620–631.
[53] C.O. Ancuti, C. Ancuti, R. Timofte, C. De 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.
[54] C.O. Ancuti, C. Ancuti, M. Sbert, R. Timofte, Dense-haze: a benchmark for image dehazing with dense-haze and haze-free images, in: 2019 IEEE International Conference on Image Processing (ICIP), IEEE, 2019, pp. 1014–1018.
[55] M.J. Islam, Y. Xia, J. Sattar, Fast underwater image enhancement for improved visual perception, IEEE Robot. Autom. Lett. 5 (2) (2020) 3227–3234.
[56] L. Peng, C. Zhu, L. Bian, U-shape transformer for underwater image enhancement, IEEE Trans. Image Process. 32 (2023) 3066–3079, https://doi.org/10.1109/TIP.2023.3276332.
[57] P. Sharma, I. Bisht, A. Sur, Wavelength-based attributed deep neural network for underwater image restoration, ACM Trans. Multimed. Comput. Commun. Appl. 19 (1) (2023) 1–23.