Image Defogging Algorithm Based on Attention Mechanism and Split Convolution

Yuanbin Wang, Yu Duan, Huaying Wu

Abstract—**In foggy days, the color saturation and contrast of the image are reduced due to factors such as scattering and refraction of light. Aiming at the problems of fog residue and loss of detailed features after processed by the existing dehazing methods, a defogging model network APSA-DehazeNet (Adaptive Pyramid Split Attention-DehazeNet) based on split convolution is proposed in this paper. Firstly, the adaptive multi-scale feature fusion module is used to extract the features of the foggy image, capture the features of different scales, then perform weighted fusion and obtain the shallow features of the image. Secondly, the deep features of the image are further obtained by the split convolutional network (PSANet) based on attention mechanism, and a more thorough dehazing effect is obtained. Finally, to effectively solve the problem of detail loss caused by a single loss, a joint loss function of perceived loss and structural similarity loss is proposed. Compared with other algorithms, experimental results demonstrate that the PSNR and SSIM indexes on the synthetic fog map dataset are improved by an average of 4.7dB and 7.4%, respectively.**

Index Terms—**image processing, image defogging, attention mechanism, multiscale networks, split convolution**

I. Introduction

IN foggy weather, numerous water molecules and suspended particles exist in the atmosphere. These elements can potentially disrupt the direct path of light, leading to light reflection from objects being diminished. The contrast and definition of images obtained by observers are easy to change, and a large number of details are lost [1-2]. These low-quality images will affect other advanced vision tasks, such as target detection, classification, tracking, traffic monitoring, and intelligent navigation [3], and bring obstacles and challenges to computer vision systems. Therefore, image defogging has important practical meaning [4].

In recent years, there have been many defogging algorithms for single images, which can be divided into two categories: traditional defogging algorithms and deep learning-based defogging algorithms. The traditional defogging algorithm mainly utilizes the imaging principle of fog, the scattering, and attenuation of light [5-6], and then establishes a model. The atmospheric scattering model consists primarily of two modules: incident light attenuation and atmospheric light imaging. By analyzing the captured image light, the process of obtaining the image is modeled according to the scattering effect of light and atmospheric optical model, and then the defogged image is obtained. Among them, the classic image defogging algorithm based on dark channel prior was proposed by HE [7] in 2009. However, the haze in the sky cannot be dealt with in this algorithm. Jin [8] proposed a defogging algorithm based on guided filtering and adaptive tolerance, which, to some extent, solved the problem that the dark channel prior algorithm distorted the sky, and subsequent algorithms improved dark channel priori (DCP) according to this point [9-11]. Although the traditional defogging algorithms have achieved certain results, there are still some limitations due to over-reliance on prior knowledge, for example, the defogging effect on the image sky domain is poor.

As deep learning advances and comprehensive defogging datasets emerge, researchers have started employing deep learning techniques for single-image defogging. At first, the researchers estimated the unknown parameters in the atmospheric scattering model by constructing a neural network and then used the estimated parameters to defog the image. The DehazeNet defogging model proposed by Cai [12] estimated the transmission map in the atmospheric scattering model with high accuracy, but only a single light source was considered. The defogging effect on outdoor pictures was poor, and the haze in the sky area was difficult to remove completely. Ren [13] proposed the Multi-scale Deep Fog Removal Network (MSCNN), utilizing a multi-scale neural network to estimate the transmission map progressively. It was not limited by the scene and had better generalization. However, it did not learn the atmospheric light value, so it had a poor effect on night smog image processing. Li [14] modified the atmospheric scattering model by merging the atmospheric light value and transmission map into a single parameter. They designed a lightweight All-in-One Network (AOD-Net) with multi-scale integration, which reduced the calculation error of the two parameters. However, due to its simple network structure, it failed to extract the image features deeply, resulting in low brightness and defogging of the defogged image. To solve this problem, a method of estimating fog-free images based on deep convolution neural network is proposed directly or iteratively. These methods mainly use the general network architecture to estimate the transmission map directly, atmospheric light value, and fog-free image, and improve the defogging performance on

Manuscript received January 7, 2023; revised August 17, 2023. This work is supported by the National Natural Science Foundation of China and National Key Research (Grants 52174198), Shaanxi Qinchuangyuan (2022KXJ-166), and the National key research and development program of Shaanxi Province, China (2023 YBSF-133).

Yuanbin Wang is an Associate Professor of Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an 710054, China (e-mail: wangyb998@163.com).

Yu Duan is a postgraduate student in Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an 710054, China (e-mail: 307073082@qq.com).

Huaying Wu is a postgraduate student in Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an 710054, China (e-mail: 1661349504@qq.com).

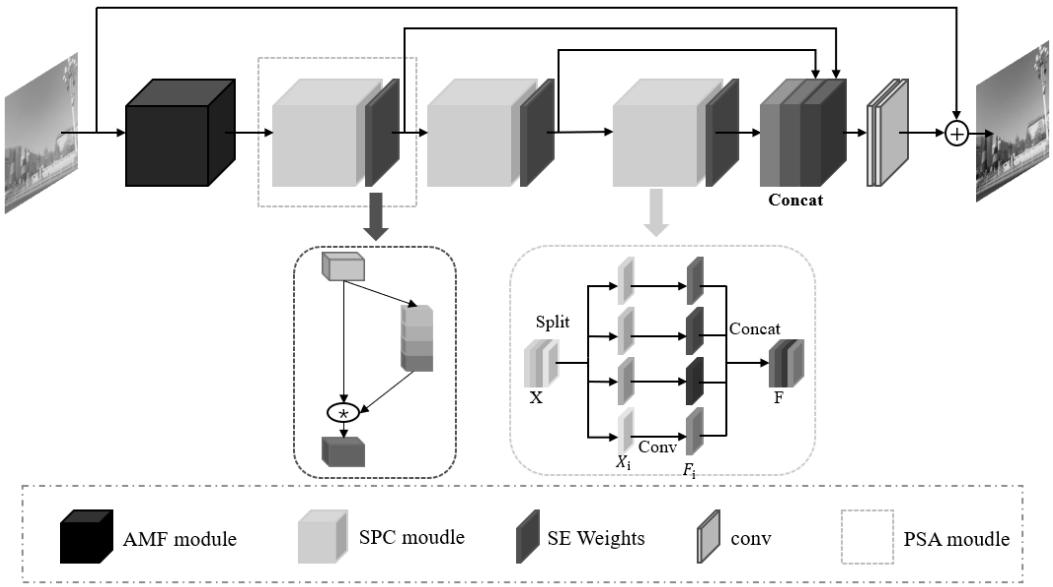

Fig. 1. The framework of the network

the premise of ensuring robustness. Chen [15] proposed a Gated Context Aggregation Network (GCANet) for direct restoration of the clear image from fog. This network utilized the latest smooth dilation technique to eliminate mesh artifacts from dilation convolution and incorporated a gated sub-network to merge features from various levels. Although the above algorithms have improved the defogging performance to a certain extent, these algorithms deal with the channel features and pixel features of the foggy image on average, resulting in insufficient attention to the dense fog pixel areas and important channel information in the image, which ultimately affects the defogging performance.

The existing defogging methods based on deep learning have exhibited promising results on publicly available datasets. However, these approaches are limited in their ability to consider local information while neglecting to establish long-range dependencies solely. The attention mechanism enhances the feature representation [16] by focusing on the important parts of the image, such as the foggy areas, which can effectively improve the defogging performance. Additionally, the networks fail to capture information at multiple scales, resulting in the inadequate enrichment of the feature space, thereby leading to issues such as the loss of fine details and incomplete defogging. To address these issues, this paper proposes an end-to-end image defogging network that utilizes patch convolution and multi-scale feature fusion. Firstly, the multi-scale network structure is utilized to extract features of different scales, and then detailed features are obtained by split convolution [17]. Ultimately, the attention mechanism is incorporated to assign weights to individual channels, facilitating the fusion of context information from various scales. This process enhances pixel-level attention, resulting in restored images with improved detailed features and defogging outcomes.

II. Network Structure

In this paper, PSA-DehazeNet, an end-to-end trainable defogging algorithm, is designed. This method is mainly composed of two parts, namely, the Adaptive Multi-scale Feature Fusion (AMF) module and the Pyramid Split Attention (PSA) module, in which the PSA module includes the Split Convolution (SPC) module and the attention mechanism. Firstly, the foggy image passes through the AMF module to capture information from the input image at different scales. Then, the split convolution and attention mechanism are used to introduce detailed features of the image. The backbone network, consisting of three PSA modules, is trained and learns to extract foggy features from the image, ultimately outputting the fog-free image. Figure 1 illustrates the framework based on split convolution and multi-scale feature fusion networks.

A. Adaptive multi-scale feature fusion

Although standard parallel multi-scale networks can effectively fuse different feature layers, their essence is simply adding different features. However, due to the inconsistency of fog features in different images, features with different resolution sizes are produced during training. If these features are simply added together in parallel multi-scale networks, features of different sizes of the same type will give unequal weights to the fused features. Large-sized features will have a greater influence on the network, while small-sized features will contribute less, significantly affecting the effectiveness of the final training model. To address this issue, an adaptive multi-scale feature fusion module is proposed, referring to the Adaptive Spatial Feature Fusion Network [18]. As depicted in Fig. 2, this module increases the weight of fused features with different sizes, dynamically and learnably adjusts the contribution of each scale, allowing the network to integrate features of different sizes better. At the same time, residual connections are added to enhance the expression ability of features, resulting in improved fog features.

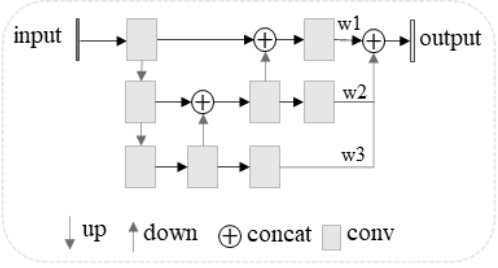

Fig. 2. Moudle of Adaptive multi-scale feature fusion

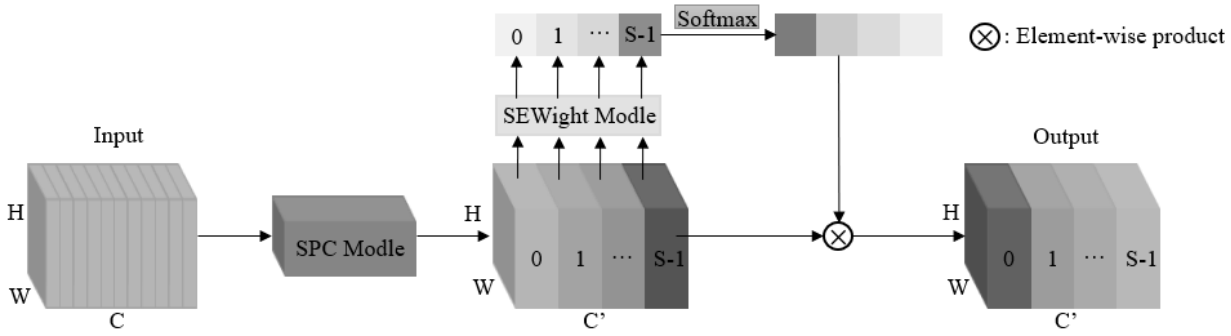

Fig. 3. Moudle of pyramid split attention

The concrete implementation is as follows: Firstly, the input features are downsampled to obtain feature maps of different scales, forming feature pyramids. Convolution operations are then applied to extract feature information from each scale. Subsequently, the output of each feature layer is upsampled to generate feature maps containing different scales. Finally, weights are assigned to the feature maps to enable adaptive feature fusion. This allows the network to objectively distribute the contribution degree of each scale to the fog features, resulting in finer fog features.

B. Pyramid split attention

In traditional defogging methods, the impact of different spatial feature maps on image details is often overlooked, leading to the loss of certain detailed features when dealing with images with inconsistent fog distribution. Therefore, we propose a PSA module based on the attention mechanism and split convolution for feature extraction in foggy images. As illustrated in Fig. 3, it is primarily implemented in four steps. Firstly, the proposed SPC module is employed to obtain multi-scale feature maps of channels. Secondly, the SEWeight module is employed to extract attention from feature maps of varying scales, leading to the acquisition of a channel attention vector. Thirdly, the recalibration weight of the multi-scale channel is computed by recalibrating the channel attention vector with Softmax. Finally, the element product operation is applied to recalibrated weights and corresponding feature maps. Through these four steps, the PSA module effectively preserves the detailed features of the image and produces a detailed feature map with richer multi-scale feature information.

1) Channel attention mechanism

With the introduction of attention mechanism, more attention is given to the fog features in the image, resulting in an improved defogging effect. The channel attention mechanism enables the network to selectively assign importance to each channel, resulting in more informative outputs. Let $x \in r$, a tensor with the size of $C \times H \times W$, represents the input feature graph, where H, W, and C represent its height, width, and the number of input channels, respectively. The SE attention mechanism [19] consists of two components: squeezing and excitation. These components are employed to capture global information and dynamically adjust the interactions among channels. Generally, channel statistics are derived through global average pooling, which integrates global spatial information into channel descriptors. The global average pooling operator for the C channel can be computed using Formula (1):

$$g_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_c(i, j) \tag{1}$$

g_c represents the global average weight, and x_c denotes the characteristic map of the channel in which it is located.

The attention weight of channel c in the SE block is represented by formula (2):

$$w_c = \sigma\left(W_1 \delta\left(W_0\left(g_c\right)\right)\right) \tag{2}$$

The symbol δ represents the Rectified Linear Unit (ReLU) operation, and W_0 and W_1 denote the fully connected layers. By utilizing two fully connected layers, the linear relationships between channels can be better integrated, facilitating the interaction between high-dimensional and low-dimensional channel information. The excitation function, denoted by σ, is commonly represented by the Sigmoid function in practical applications. Through the implementation of the excitation function, channel weights are assigned after channel interaction, enabling enhanced information extraction. This module is denoted as the SEWeight module, and its visual representation is illustrated in Figure 4.

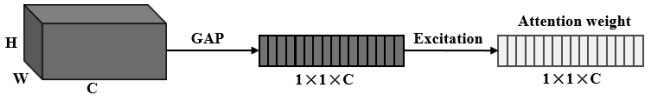

Fig. 4. Moudle of SEWeigh

2) Split convolution module

The SPC is the fundamental operator for multi-scale feature extraction in the PSA module. The input feature map, denoted as X, is divided into S parts along the channel dimension, represented as X_1, X_2, ..., X_{S-1}. Each divided part has S shared channels, and the resulting feature map is denoted as X_i. C is divisible by S, where C represents the number of channels in the input feature map.

Next, we process the input feature maps of multiple scales in parallel using split convolution, which enables us to obtain feature maps that contain individual kernel types. This method enables the extraction of spatial information from each channel feature map and generates varying spatial resolutions and depths using a pyramid structure with multi-scale convolution kernels. It learns multi-scale spatial information and establishes local cross-channel interaction independently for each segmented part.

At the same time, we design a standard to select the group size. Formula (3) describes the correlation between the multi-scale kernel size and the group size:

$$G = 2^{\frac{K-1}{2}} \tag{3}$$

Where K represents the convolution kernel size, and G stands for the group size. The formula (4) for generating the multi-scale feature map is as follows:

$$F_i = Conv\left(k_i \times k_i, G_i\right) X_i, i = 0,1,2\mathrm{L}\ \ S-1 \tag{4}$$

Where the k_i denotes kernel size, and the ith group sizes G_i and F_i represent feature graphs of different scales. The comprehensive multi-scale feature map is achieved by concatenating the individual feature maps. By deriving channel attention weights from the preprocessed multi-scale feature map, attention weight vectors for each scale are acquired. Mathematically, the vector of attention weights can be expressed using formula (5):

$$Z_i = SEWeight\left(F_i\right), i = 0,1,2\mathrm{L}\ \ S-1 \tag{5}$$

Where Z_i represent the attention weight. Attention weights of different scales are obtained by SEWeight module from input feature maps. The PSA module fuses context information of various scales, enhancing pixel-level attention for advanced feature maps. Finally, the weighted weights are normalized by the softmax function and multiplied by the output of the SPC module to obtain the feature output with multi-scale information.

3) Loss function

L_P loss is the perceptual loss [20-21], which is the loss function of the model. It compares the features obtained through convolution of actual images with those obtained through convolution of generated images, aiming to align high-level information (content and global structure), thus achieving perceptual similarity. Recently, researchers have discovered that utilizing perceptual loss through feature comparison better aligns with human visual perception and preserves finer details. Compared with the mean square error loss, the perceptual loss changes the computational space from image space to feature space, and the computational formula is shown in formula (6):

$$L_p = \frac{1}{C_j H_j W_j}\left\|\varphi_j\left(y\right) - \varphi_j\left(\dot{y}\right)\right\|_2^2 \tag{6}$$

Where φ represents the characteristic map of VGG16, y is a clear image and $\dot{y}$ is a defogged image.

The function of image similarity loss function is to make the image visual effect more in line with the subjective visual feeling of human eyes. It can intuitively reflect the degree of structural similarity between the generated image and the standard clear image [16]. The loss function can be expressed as shown in the formula (7):

$$L_S = 1 - \frac{\left(2\mu_x\mu_y + c_1\right)\left(2\sigma_{xy} + c_2\right)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)} \tag{7}$$

Where x, y represents defogged image and clear image, respectively, and μ and σ are the mean and variance of images. σ_{xy} is covariance, c_1 and c_2 are variables.

To sum up, the overall loss function of this method is shown in Formula (8):

$$L_{total} = \alpha L_P + \beta L_S \tag{8}$$

In the formula, α and β represent the weight hyperparameters used to control L_P and L_S losses, respectively. We measured the magnitude and convergence speed of the two losses during the experiment, and set α and β to 0.95 and 0.05, respectively.

III. Experimental analysis

A. Dataset and experimental settings

We utilized the RESIDE dataset as a benchmark for image dehazing tasks in the computer vision community. The training set comprised the outdoor (OTS) and indoor (ITS) datasets in the RESIDE dataset, containing 13,990 synthetic blurred images generated from 1,399 clear images under various atmospheric scattering conditions. These images are synthesized with different values of atmospheric scattering coefficients A ∈ [0.7,1.0] and transmission coefficients β ∈ [0.6,1.8] based on the atmospheric scattering model. To assess the effectiveness of our proposed method, we selected the comprehensive target dataset (STOS) as the test set, which comprises 200 real-world outdoor images captured under diverse weather and lighting conditions. We compared our method with several state-of-the-art dehazing algorithms on this test set, including DCP, DehazeNet, AOD-Net, and GCANet.

In this paper, a supervised learning method is adopted, and the sizes of the input synthetic foggy images and the corresponding original foggy images are both modified to 640×480 RGB three-channel images. The batch size is 16, the learning rate is set to 0.0001, and the total number of iterations is 1000. The Adam optimizer is used for optimization, and the weight decay parameter is set to 0.0001. The entire experiment is carried out by training on an NVIDIA 3070 GPU.

Peak Signal-to-Noise Ratio (PSNR) is a widely used metric to assess image quality. It measures the ratio of the maximum possible power of a signal to the power of corrupting noise, and higher PSNR values indicate better defogging performance by indicating lower levels of noise. The Structural Similarity Index (SSIM) is another evaluation metric that combines multiple components such as brightness, contrast, and structure to assess the similarity between two images. A higher SSIM value signifies a closer level of similarity between the defogged image and the original image. Therefore, PSNR and SSIM are chosen as the objective evaluation criteria for assessing the defogging algorithm's performance.

B. Indoor image defogging

Fig. 5 compares the results of each algorithm on an indoor synthetic fog map. The DCP algorithm shows a good defogging effect, but also reduces contrast and darkens the colors. The DehazeNet algorithm over-enhances the image, resulting in the loss of image details. The AOD-Net algorithm still leaves fog residue on the entire image. Both GCANet and the proposed algorithm effectively defog indoor images, but there is still a slight color distortion observed in the images processed by GCANet.

Table I shows the average defogging performance of the five algorithms on indoor fog maps. The traditional DCP algorithm has low PSNR and SSIM values. DehazeNet, a

TABLE I
PERFORMANCE OF DIFFERENT ALGORITHMS ON INTDOOR FOG MAPS

	DCP	DehazeNet	AOD-Net	GCANet	Proposed
PSNR	18.386	17.621	20.400	22.331	**22.852**
SSIM	0.832	0.754	0.868	0.923	**0.934**

Fig. 5. Comparison of ITS indoor image

deep learning-based algorithm, has the poorest defogging effect on indoor images, resulting in the lowest evaluation index. Although the evaluation indexes of AOD-Net and GCANet have improved, they do not take into account the difference in fog distribution between channels. In contrast, the proposed algorithm in this paper pays attention to the fog characteristics between different channels by using split convolution and attention mechanism. It assigns weights to different channels so that features of different scales can be reasonably utilized, resulting in a more thorough fog removal effect. As a result, the proposed algorithm achieves better evaluation indexes, with PSNR and SSIM values 0.521dB and 0.011 higher than that of GCANet, respectively.

C. Outdoor image defogging

Fig. 5 shows the comparison of defogging results of each algorithm on an outdoor fog map. It can be observed from the figure that the DCP algorithm causes serious color distortion after defogging. DehazeNet over-enhances the image, resulting in residual fog in the distant areas of the image. AOD-Net decreases the overall contrast of the image, and the defogging effect is incomplete. GCANet has an obvious defogging effect, but there is still color distortion in the sky area and distant parts of the image, and the color is darker than in the original image. Compared with the above-mentioned algorithms, proposed algorithm exhibits a better visual effect.

TABLE II
PERFORMANCE OF DIFFERENT ALGORITHMS ON OUTDOOR FOG MAPS

	DCP	DehazeNet	AOD-Net	GCANet	Proposed
PSNR	15.363	20.187	19.845	22.416	**23.614**
SSIM	0.711	0.862	0.829	0.914	**0.933**

Table II presents the defogging evaluation indexes of each algorithm on outdoor fog map. It is evident that the proposed algorithm has the best evaluation index, with a PSNR value and SSIM value of 23.614 and 0.933, respectively. In contrast, the PSNR value of the GCANet algorithm is about 22

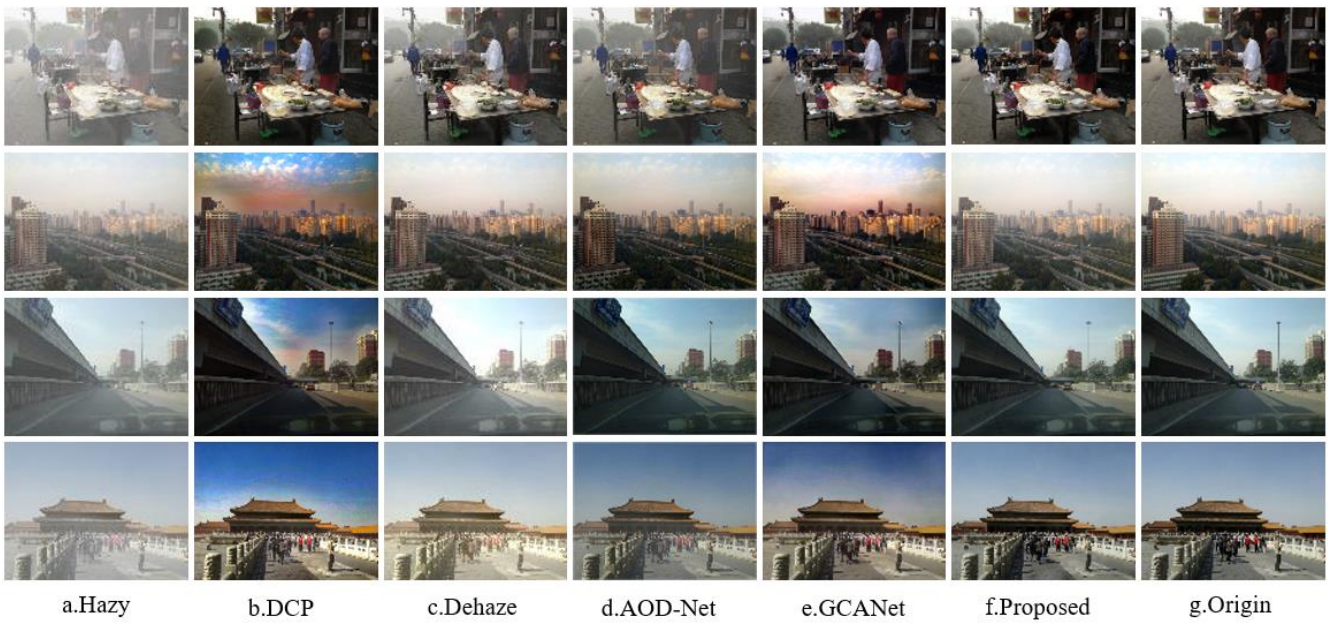

Fig. 5. Comparison of ITS indoor image

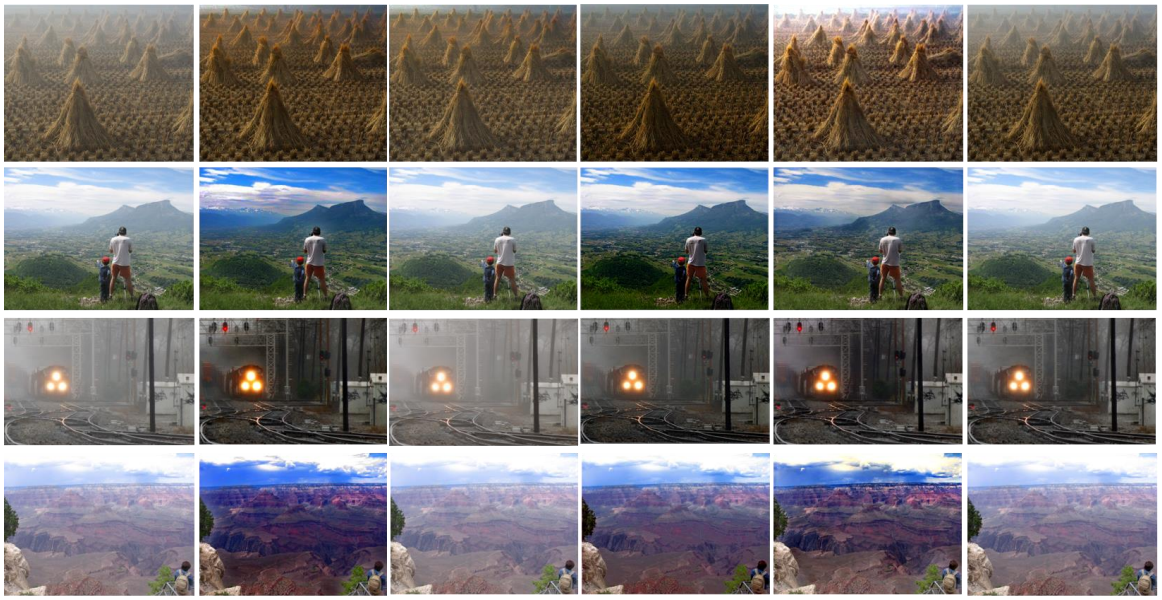

Fig. 6. Comparison of real scene

while the values of AOD-Net and DehazeNet algorithms are similar. The images in the SSIM groups are uneven, and both GCANet and the proposed algorithm perform well. By utilizing the adaptive multi-scale feature fusion module, this paper obtains richer image details and reduces the loss of details. Therefore, the proposed algorithm has better defogging performance and a more natural visual effect compared to other algorithms.

D.Real hazy image defogging

Figure 6 displays the results of applying the proposed algorithm to authentic fog images in order to verify its dehazing effect. Observing the figure, it becomes apparent that the DCP algorithm leads to significant color distortion in the image. The DehazeNet algorithm still leaves a considerable amount of fog residue in the image, while the AOD algorithm causes a significant reduction in image contrast. On the other hand, GCANet produces grid artifacts in the central portion of the defogged image. In comparison to these algorithms, the proposed algorithm better preserves the color of the image while effectively reducing the fog.

For the real scene, due to the lack of original clear image as a comparison, this paper uses two objective indicators of no-reference evaluation index information entropy (IE) and average gradient (AG) to evaluate the dehazing effect of the real fog map. IE can be used as an index to measure the amount of information in the image, the larger the IE value, the better the defogging effect. AG can reflect the details and texture information in the image, the larger the AG value, the more details appear in the image, the clearer the image.

In Table III, the objective indicators of each algorithm on the real fog map are presented. The proposed method shows the best performance in terms of information entropy, indicating that it effectively reduces the amount of fog and enhances the clarity of the image. However, when it comes to the average gradient index, which measures edge details and texture information, the proposed algorithm performs relatively poorly. To preserve the original color of the image, this algorithm results in a slight compromise in the definition and detail of the edge.

TABLE III
PERFORMANCE OF DIFFERENT ALGORITHMS ON REAL SCENE

	DCP	DehazeNet	AOD-Net	GCANet	Proposed
IE	7.210	7.110	6.636	7.512	**7.781**
AG	7.422	6.768	**8.125**	7.941	7.535

E. Ablation experiment

To evaluate the effectiveness of the proposed loss function for image defogging, the model is trained using different weightings of the loss function, and the resulting defogging effects are analyzed.

TABLE IV
EXPERIMENTAL COMPARISON OF DIFFERENT LOSS FUNCTION WEIGHTS

Weight	Loss function	PSNR	SSIM
α=1, β=0	L_P	19.289	0.884
α=0.85, β=0.05	0.85 * L_P + 0.05 * L_S	20.664	0.907
α=0.90, β=0.05	0.90 * L_P + 0.05 * L_S	21.463	0.913
α=0.95, β=0.05	0.95 * L_P + 0.05 * L_S	**22.524**	**0.928**
α=0.95, β=0.10	0.95 * L_P + 0.10 * L_S	22.011	0.916

In Table IV, the dehazing effects of various loss function weights are compared. It is observed that when using only the perceptual loss function, the objective evaluation index of the defogged image is low. This is because the perceptual loss function may not fully capture the spatial similarity of the image. However, incorporating the structural similarity loss function into the model led to a significant improvement in the performance index of the defogged image. Specifically, the experimental results indicated that the best objective evaluation index of the image is obtained when the loss function is weighted with α=0.95 and β=0.05.

TABLE V
ABLATION EXPERIMENTS OF EACH MODULE

	AFM	PSA	PSNR/dB	SSIM
Model 1	×	×	20.387	0.892
Model 2	√	×	21.473	0.908
Model 3	×	√	21.722	0.917
Model 4	√	√	**22.542**	**0.928**

In order to verify the effectiveness of the module proposed in this paper, experiments are conducted on the SOTS dataset. The results of four experimental indexes are shown in Table V. It can be observed that the PSNR and SSIM of Model 2 and Model 3 improved significantly compared to those of the basic defogging network. This indicates that the AFM module and PSA module can improve the ability to extract and reconstruct fog feature information, thus retaining more details and making the restored image more complete, resulting in a better defogging effect.

IV. Conclusion

In this paper, we propose an image defogging network based on split convolution and multi-scale feature fusion. The proposed network is evaluated on indoor and outdoor scenes, demonstrating its effectiveness for subsequent vision tasks. The network utilizes an adaptive multi-scale feature fusion structure to extract foggy image features, resulting in better feature extraction and utilization. Furthermore, we introduce a PSA module that leverages split convolution to extract fog features in-depth, capturing and fusing information between channels using an attention mechanism. This results in a more thorough fog removal effect. Experimental results demonstrate that the proposed algorithm outperforms common defogging algorithms, effectively removing fog while preserving detailed information and achieving a natural visual effect.

While the proposed network has shown promising results, there are several limitations that need to be addressed in future research. The effectiveness of the proposed network is contingent upon supervised learning and a substantial volume of training data. The network's performance could be constrained by the adequacy and quality of available training data. Therefore, it will be necessary to explore unsupervised or weakly supervised learning methods to reduce the reliance on annotated data. Moreover, the proposed network focuses on removing uniform fog but may need to perform better for non-uniform fog or haze. Developing a network capable of handling various types of haze is another important direction for future work.

References

[1] J. PENG, F. J. XUE and Y. B. YUAN, “Adaptive Image Defogging Algorithm Combining Multi-Scale Retinex and Dark Channel,” [J]. *Laser & Optoelectronics Progress*, vol. 58, no. 04, pp. 117–125, 2021.

[2] R. KUMAR, B. K. KAUSHIK and R. BALASUBRAMANIAN, “Multispectral transmission map fusion method and architecture for image dehazing,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 27, no. 11, pp. 2693–2697, 2019.

[3] H. ZHANG and V. M. PATEL, “Densely connected pyramid dehazing network,” *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3194-3203, 2018.

[4] D. WU and Q. S. ZHU, “Recent research advances in image defogging,” *Acta Automatica Sinica*, no. 02, pp. 221-239, 2015.

[5] L. P. YAO, and Z. L. PAN, “The Retinex-Based Image Dehazing Using a Particle Swarm Optimization Method,” *Multimedia Tools and Applications*, vol. 80, no.3, pp. 3425–42, 2020.

[6] P. Y. LI, J. D. TIAN, and Y. D. TANG, “Deep Retinex Network for Single Image Dehazing,” *IEEE Transactions on Image Processing*, vol. 30, no.12, pp. 1100–1115, 2020.

[7] 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*, vol. 33, no.12, pp. 2341-2353, 2010.

[8] Q. Tang, J. Yang and X. He, “Nighttime image dehazing based on Retinex and dark channel prior using Taylor series expansion,” *Computer Vision and Image Understanding*, pp. 103086, 2021.

[9] T. Y. GUO, N. LI and L. L. SUN, “A Defogging Method Containing Images of Sky,” *Laser & Optoelectronics Progress*, val. 58, no.20, pp.2010011, 2021.

[10] A. GOLTS, D. FREEDMAN and M. ELAD, “Unsupervised single image dehazing using dark channel prior loss,” *IEEE Transactions on Image Processing*, no. 29, pp. 2692-2701, 2019.

[11] H. F. Pu, Z. Li and L. R. Li, “Research of Dehazing Algorithm Based on Dark Channel Prior,” *Laser & Optoelectronics Progress*, val. 58, no.20, pp.2010011, 2021.

[12] B. CAI, X. XU and K. JIA, “DehazeNet: An end-to-end system for single image haze removal,” *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5187-5198, 2016.

[13] W. Q. REN, S. LIU and H. ZHANG, “Single image dehazing via multi-scale convolutional neural networks,” *European Conference on Computer Vision. Amsterdam*, pp. 154-169, 2016.

[14] B. Y. LI, X. L. PENG and Z. Y. WANG, “Aodnet: all-in-one dehazing network,” *IEEE International Conference on Computer Vision*, pp. 4780-4788, 2017.

[15] D. D. CHEN, M. M. HE and Q. N. FAN, “Gated context aggregation network for image dehazing and deraining,” *IEEE Winter Conference on Applications of Computer Vision*, pp. 1375-1383, 2019.

[16] C. CHEN, B. WU, and H. Y. ZHANG, "An Image Recognition Technology Based on Deformable and CBAM Convolution Resnet50," *IAENG International Journal of Computer Science*, vol. 50, no.1, pp274-281, 2023.

[17] S. LIU, D. HUANG and Y. H. WANG, “Learning Spatial Fusion for Single-Shot Object Detection,” *arXiv: Computer Vision and Pattern Recognition*, 2019.

[18] H. ZHANG, K. K. ZU and J. LU, “EPSANet: An Efficient Pyramid Split Attention Block on Convolutional Neural Network,” *arXiv: Computer Vision and Pattern Recognition*, 2021.

[19] H. JIE, S. LI and S. GANG, “Squeeze-and-excitation networks,” *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[20] J. JUSTIN, A. ALEXANDRE and F. F. LI, “Perceptual Losses for Real-Time Style Transfer and Super-Resolution,” *In Computer Vision – ECCV 2016, Lecture Notes in Computer Scienc*e, 2016.

[21] M. S. RAD, B. BOZORGTABAR and U. V. MARTI, “Srobb: Targeted perceptual loss for single image super-resolution,” *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019.