Research on Image Defogging Algorithm Based on Improved FFA-Net

Li Qinrong, Ma Chi*, Guo Qiang, Hu Hui

Abstract—**Images captured under severe weather conditions, such as haze and fog, suffer from image quality degradation caused by atmospheric particle diffusion. This degradation manifests as color fading, reduced contrast, and adversely affects the performance of various computer vision tasks. To address this, this paper presents an end-to-end feature fusion attention network (FFA-Net) designed to directly restore haze-free images. By incorporating the SSIM loss into the original loss function, the proposed method effectively captures the visual disparities between the estimated defogged image and the authentic haze-free image. Additionally, it mitigates the color distortion problem inherent in the original algorithm.**

To address the challenge of low brightness in input images, a low illumination enhancement module is introduced, seamlessly integrated with the FFA-Net defogging method. Subsequently, a comparative analysis of different defogging algorithms is conducted using two distinct foggy datasets. Multiple evaluation metrics are employed to assess the performance of these algorithms. The findings indicate that our algorithm significantly outperforms others in terms of objective indicators such as PSNR and SSIM, as well as visual effects.

Index Terms—**Deep learning, Image defogging, FFA-Net, Low illumination enhancement module**

I. Introduction

IMAGES obtained under bad weather conditions such as haze and fog, due to the diffusion of atmospheric particles, deteriorate the image quality, resulting in color fading and reduced contrast. At the same time, the texture and edges of objects in the scene become blurred, which poses a challenge to human interpretation and target feature recognition [1]. Many researchers are striving to enhance the quality of hazy images to produce clearer scenes, which holds significant practical importance for people's daily lives and productivity. Image defogging techniques can be categorized into three types based on their fundamental principles and features: methods focused on image enhancement [2], those centered on restoration [3], and those employing deep learning [4].

Manuscript received November 6, 2023; revised April 19, 2024.

This paper is supported by Foundation on Guangdong Educational Committee under the Grant No. 2022ZDZX4052, No. 2021ZDJS082.

Li Qinrong is a postgraduate student of School of Computer Science and Software Engineering, University of Science and Technology LiaoNing. AnShan 114051, China (e-mail: 1321380131@qq.com).

Ma Chi is an asssociate professor of the School of Computer Science and Engineering, Huizhou University. Huizhou 516007, China (*corresponding author to provide e-mail:machi@hzu.edu.cn).

Guo Qiang is a postgraduate student of School of Computer Science and Software Engineering, University of Science and Technology LiaoNing. AnShan 114051, China (e-mail: 254755139@qq.com).

Hu Hui is an asssociate professor of the School of Computer Science and Engineering, Huizhou University. Huizhou 516007, China (e-mail:huhui@hzu.edu.cn).

The primary objective of image enhancement-based defogging methods is to eliminate noise from images while simultaneously improving the captured information to produce high-quality results. To achieve this goal, there are three widely used techniques: histogram equalization [5], Retinex-based algorithms [6], and frequency-based image enhancement algorithms [7]. Histogram equalization is a method that doesn't directly address the physical degradation of images, but instead, it uniformly adjusts gray levels across a particular spectrum to enhance image contrast. In contrast, the Retinex algorithm, which is based on the retinal cerebral cortex theory proposed by Land and McCann [8], takes into account the human eye's color perception and improves image quality by enhancing color constancy. Frequency-based image enhancement algorithms concentrate on enhancing blurry images that have low-frequency characteristics in the frequency domain. These algorithms introduce a high-pass filter during image filtering to compensate for low frequencies, thereby enhancing high-frequency details.

The restoration-based method involves analyzing the imaging mechanism and degradation specifics of an image, then employing inverse transformations to refine the degraded image into high-quality versions [9]. Many researchers in this field generalize this process into a physical model for image restoration based on image degradation [10]. McCartney [11] proposed an atmospheric physics model, establishing atmospheric light scattering and light attenuation models. Narasimhan's atmospheric model [12], based on McCartney's method, suggests that the overall radiant intensity recorded by the camera is associated with the linear combination of scene radiation, encompassing scattered light that enters the capturing device. Tan and Oakley [13] expanded Oakley's algorithm by examining the degradation model based on multi-parameter statistics, assuming knowledge of scene depth and projecting haze image restoration to color images. He, Sun, and Tang [14] introduced the Dark Channel Prior (DCP) defogging method, which relies on prior information to produce visually clear images. However, constructing the dark channel prior is computationally expensive and slow. Sun, Xiao, and Wei [15] employed an approach that accounts for variations in scene depth, adapting the atmospheric scattering model to a monochromatic form. Using this refined model, images are segmented into sky and non-sky regions, each of which is processed independently.

Before deep learning, image defogging algorithms were widely used in computer vision tasks, mainly relying on various prior information assumptions and atmospheric scattering models. These statistical rule-based processing methods have good interpretability. However, when faced with complex reality, they may exhibit drawbacks, such as

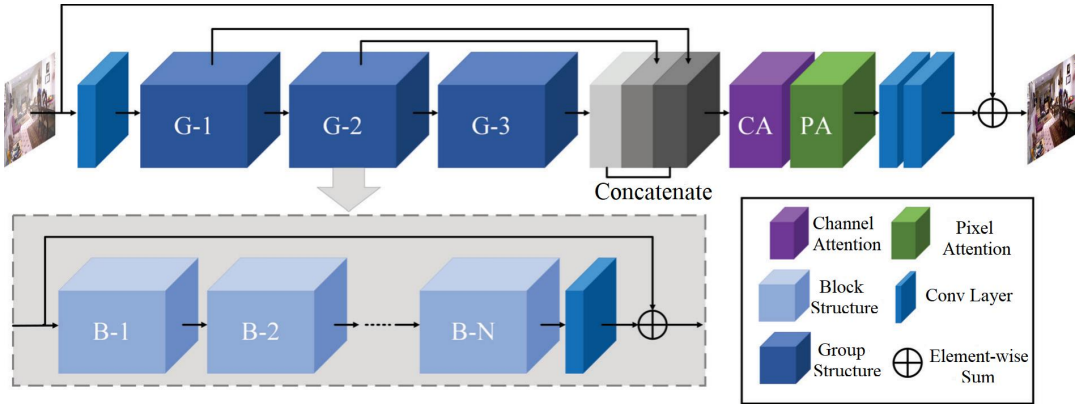


Fig. 1. FFA-Net architecture

high computational complexity and susceptibility to inter from complex environments. DehazeNet [16], inspired by deep learning, employs convolutional neural networks to estimate parameters of atmospheric scattering models. It trains on haze features and their propagation role. All-in-One Network (AOD-Net) [17] directly generates haze-free results from foggy images without separately assessing transmission and airborne light. Zhang and Patel [18] developed an end-to-end model using generative adversarial networks to remove haze, learning structural relationships from images. They also introduced an atmospheric model to enhance overall learning. Qin et al. [19] introduced an end-to-end image defogging algorithm named Feature Fusion Attention Network (FFA-Net). This algorithm utilizes channel and pixel attention modules to consider features from different channels and learn distinct features. By fusing different channel weights, it prioritizes essential features like thick haze areas.

This paper will use the FFA-Net defogging algorithm as the basic framework, and further describe the visual differences between the estimated defogging image and the real fog free image by adding SSIM loss to the original loss function, as well as alleviate the color distortion problem caused by the original algorithm. In response to the unclear image of the input image under low brightness, a low illumination enhancement module is introduced to combine with the FFA-Net defogging algorithm. Subsequently, the proposed algorithm's efficacy was demonstrated by conducting experimental comparisons with other traditional defogging algorithms using a dataset of foggy images.

II. FFA-Net Defogging Algorithm

A. FFA-Net Structure Analysis

FFA-Net is an end-to-end Feature Fusion Attention network used to restore foggy images to non foggy images. The architecture of FFA-Net is shown in Fig. 1.

Fig. 1 illustrates the working of the FFA-Net, which begins with processing a foggy and blurry image through a shallow feature extraction phase. This step creates an initial feature map. The map is then passed to sequential Group Structures, each imparting multiple skip connections. The resulting output from these Group Structures undergoes fusion via the Feature Attention (FA) module. Then, these feature information are transmitted to two convolutional layers combined with a local residual network to output the enhanced fog free image. The Block Structure comprises a local residual network and an FA module. The FA module is a structured attention mechanism comprising Channel Attention (CA) [20] and Pixel Attention (PA) [21]. Its structure is illustrated in Figure 2.

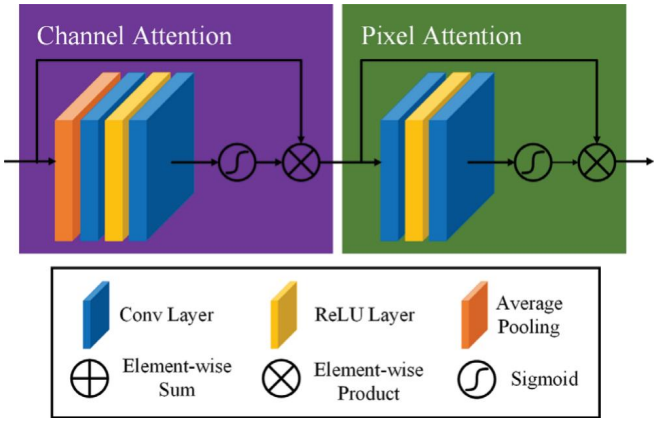


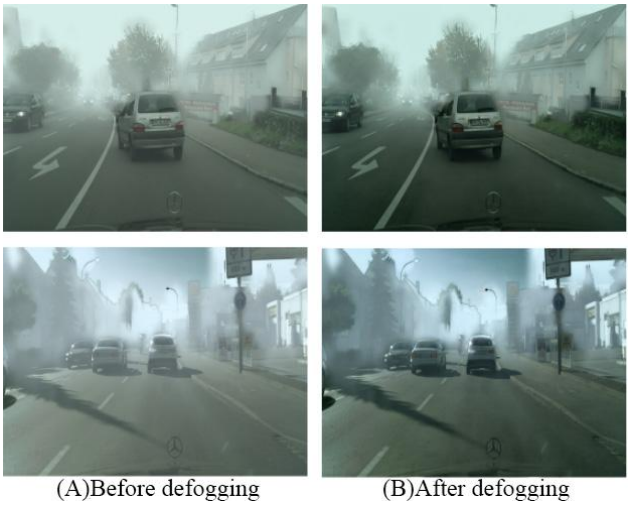
Fig. 2. FA module

Unlike most image defogging networks that do not perform weighted processing on different channels and regions of input data, which cannot effectively handle images with uneven distribution of haze concentration, FA is composed of CA and PA, which can flexibly handle different types of information. The CA module distinguishes the feature maps of different channels by weighting them, The PA module is used to weight different pixel regions of the feature map. Furthermore, each Block Structure comprises a local residual network and a FA module. The inclusion of local residual networks enhances the training stability of deep neural networks, reduces the complexity of parameter optimization, and boosts network performance by increasing its depth. A Group Structure is composed of a Block Structure and a skip connection module. The seamless design of the Block Structure significantly enhances the depth and expressive power of FFA-Net, while the skip connections prevent training issues for FFA-Net. Subsequently, the feature information can be restored to the desired clear image by utilizing a recovery model. After passing through two convolutional layers, the model concatenates the output results with the original input through residual connections.

B. FFA-Net Defogging Effect Analysis

In order to visually observe the actual defogging effect of the FFA-Net algorithm, the artificially synthesized fog

dataset Citscapes_foggy and the real fog dataset RTTS were selected as experimental subjects, and FFA-Net was used for defogging treatment on these two datasets. The performances of defogging on two datasets is shown in Fig. 3 and Fig. 4 respectively.



(A)Before defogging (B)After defogging

Fig. 3. The comparison image after dehazing on Citscapes_foggy dataset

(A) Before defogging (B) After defogging

Fig. 4. The comparison image after dehazing on RTTS dataset

From the comparison before and after defogging in Fig. 3 and Fig. 4, it can be seen that using FFA-Net for defogging can make the original haze image clearer and enhance the visual effect. This indicates that FFA-Net can restore the synthesized blurred haze image to a clear image to a certain extent, or perform defogging enhancement on real haze images. However, this algorithm mainly has the following two problems:

Firstly, as depicted in the initial set of images in Fig. 3, the algorithm's defogging effect is minimal under low light conditions. After defogging, it leads to image darkening and some degree of color distortion. Secondly, the FFA-Net algorithm exhibits poor defogging performance on real fog datasets. This is attributed to its exclusive utilization of synthetic fog datasets for training, and the narrow focus of the utilized loss function on pixel-level discrepancies, neglecting human visual system perception. Additionally, the evaluation indicators exclusively rely on reference metrics, which doesn't align with human visual habits, resulting in subpar defogging performance on actual foggy datasets.

III. Improved FFA-Net Defogging Algorithm

This section discusses methods to improve the shortcomings of the FFA-Net mentioned earlier. Initially, the original algorithm's loss function solely emphasizes pixel differences, neglecting human visual system perception. To address this, the paper introduces SSIM loss alongside the original algorithm's loss function, combining the two as the final loss function. Additionally, to tackle the influence of low lighting conditions on defogging outcomes and the problem of reduced image brightness post-defogging, a low illumination enhancement module Zero-DCE was introduced and integrated into the FFA-Net network structure.

A. Improvemnt of Loss Function

The loss function used by FFA-Net to optimize network weights only focuses on pixel level differences, without considering the perception of the human visual system. Therefore, this paper modifies the loss function of FFA-Net to enable the network to learn better pixel and visual features to improve the defogging effect. The loss functions commonly used for image defogging include the following:

L2 loss, is also known as Mean Squared Error (MSE) loss, was extensively utilized in early single image defogging and other convolutional neural network tasks. The L2 loss function has a simple expression and fast convergence speed, but L2 loss focuses more on pixel level differences and has no good correlation with the perception of the human visual system. The L2 loss function formula is as follows:

$$L_{l_2}(P)=\frac{1}{N}\sum_{p\in P}(x(p)-y(p))^2 \tag{1}$$

In this formula, P represents Patch, p is the index of the pixel, N represents the number of p in the P, $x(p)$ represents the processed pixel value, i.e. the predicted value, and $y(p)$ represents the actual value (ground truth).

L1 loss, was employed in the original FFA-Net. The advantage of the L1 loss function over the L2 loss is that it is more robust to outliers because it has lower sensitivity to outliers. This is due to the L2 loss function magnifying the loss value through the use of squared operations when calculating differences between pixels. The penalty for L1 loss in absolute value operations is smaller, less susceptible to outlier interference, more stable training, and less prone to falling into local optima. The L1 loss formula is as follows:

$$L_{l_1}(P)=\frac{1}{N}\sum_{p\in P}\left|x(p)-y(p)\right| \tag{2}$$

In this formula, the meanings of P, p, N, $x(p)$ and $y(p)$ are the same as the (1).

SSIM loss. Image defogging not only requires obtaining higher indicator scores at the pixel level, but also requires the network to learn to produce visually pleasing images. SSIM is a measure of image similarity, which comprehensively considers the similarity of brightness, contrast, and structural similarity between two images, in line with the human visual perception system. The definition of SSIM is:

$$SSIM(p)=\frac{2\mu_x\mu_y+C_1}{\mu^2{}_x+\mu^2{}_y+C_1}\cdot\frac{2\sigma_{xy}+C_2}{\sigma^2{}_x+\sigma^2{}_y+C_2} \tag{3}$$

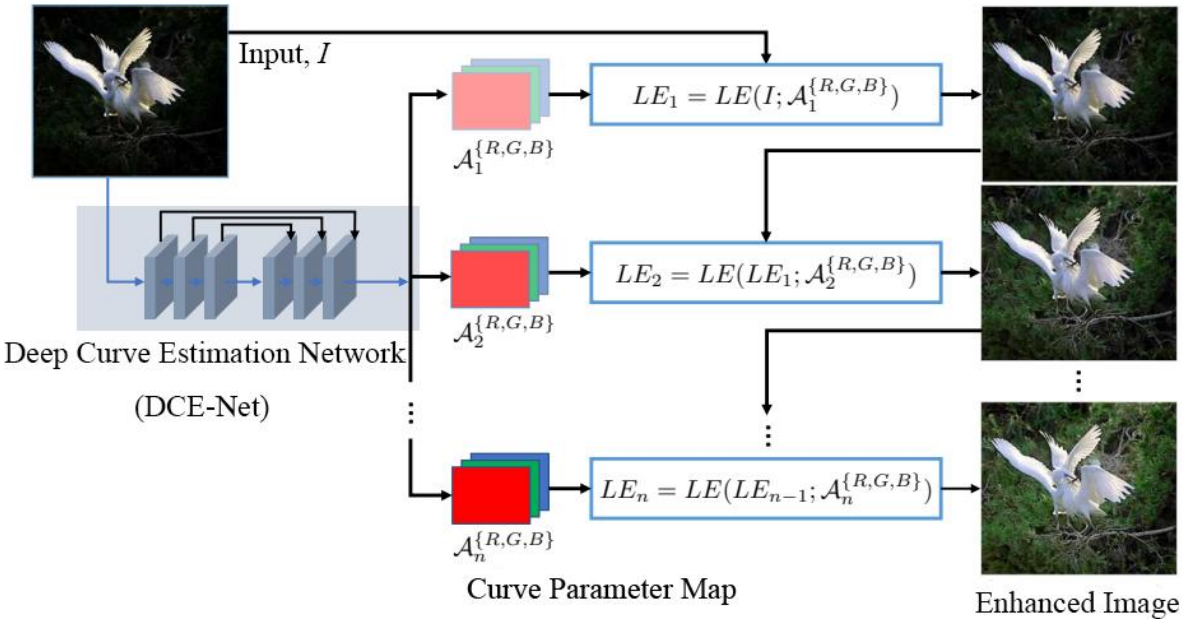


Fig. 5. The framework of Zero-DCE

In this formula, The meaning of p is the same as (1), μ_x represents the average value of x, used to estimate the brightness of x, σ_x represents the standard deviation of x, used to estimate the contrast of x. The meaning of μ_y and σ_y is similar to μ_x and σ_x. σ_{xy} representing the covariance of x and y, measuring the trend of common changes between x and y, used to estimate the structural similarity of x and y, C_1 and C_2 used to maintain a stable constant less than 1, which can avoid the phenomenon of division by 0.

The SSIM loss can be written as follows:

$$L_{SSIM}(\mathrm{P}) = \frac{1}{\mathrm{N}} \sum_{p \in P} (1 - SSIM(p)) \tag{4}$$

In this formula, the meanings of P, p, N, $x(p)$ and $y(p)$ are the same as the (1).

In order to enable the network to learn the best pixel and visual features, this paper combines L1 loss and SSIM loss, and introduces a hyperparameter to participate in network training, so that the network can obtain the best features of two error functions. The total loss function is as follows:

$$L_{total}(P) = \alpha L_{l_1}(P) + (1-\alpha) L_{SSIM}(p) \tag{5}$$

In this formula, α is utilized as a hyperparameter to adjust the relative impact of the two loss functions. Based on experimental experience, this paper sets it to 0.4, and it is found that it does not become sensitive due to small changes.

B. Low Illumination Enhancement Module

In response to the poor performance of FFA-Net under low lighting conditions and the issue of reduced image brightness after defogging, this paper introduces a low illumination enhancement module. In contrast to conventional approaches, deep learning-based methods for light enhancement can globally perceive the information of the entire image, while traditional methods are often limited to enhancing local regions. Hence, this study presents a Zero-Reference Deep Curve Estimation (Zero-DCE) [22] approach employing deep networks to improve dim images, effectively addressing diverse lighting conditions such as non-uniformity and inadequate illumination. This algorithm has the characteristics of lightweight and strong robustness. Therefore, this paper embeds it into the FFA-Net network structure to enhance low light level images.

In contrast to conventional image restoration tasks centered on image-to-image mapping, Zero-DCE employs a deep neural network to calculate high-order curves. These high-order curves are generated based on the input low-light image and are subsequently employed to adjust the dynamic range of the input pixels, leading to the production of an improved, high-brightness image. To estimate pixel dynamic range and high-order curves, Zero-DCE employs a purpose-built deep network, DCE-Net. This network comprises seven convolutional layers. A defining feature of Zero-DCE lies in its "zero reference" approach—it doesn't require reference data during the training phase. This unique attribute is facilitated by an ingenious set of non-reference loss functions. The structure of the Zero-DCE framework is presented in Fig. 5.

As shown in Fig. 5, Zero-DCE consists of three key components:

(1) Light-Enhancement Curve (LE-Curve)

The light enhancement curve automatically transforms dimly lit images into enhanced images. Light-Enhancement Curve aims to accomplish three effects: normalizing all pixel values of the enhanced image between zero and one to prevent signal loss from data truncation due to overflow, ensuring the curve is monotonically increasing to provide varying contrast with adjacent images. In order to achieve these goals, a quadratic curve model has been devised, aiming to be both straightforward and easily differentiable to facilitate enhanced gradient backpropagation. This model is described as follows:

$$LE(I(x);\alpha) = I(x) + \alpha I(x)(1 - I(x)) \tag{6}$$

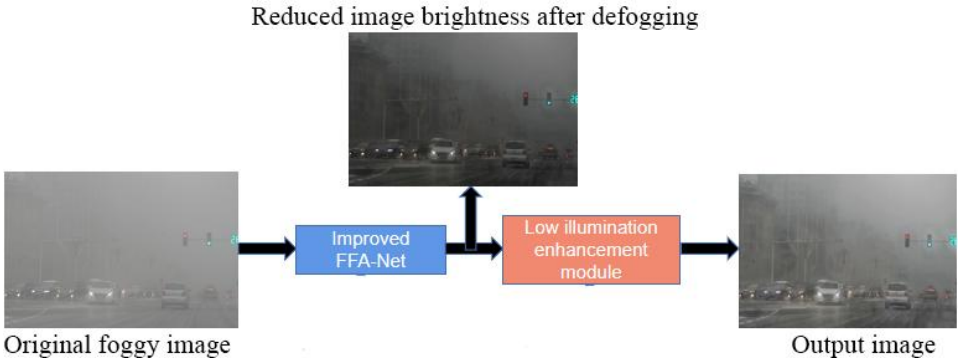


Fig. 6. The framework of improved defogging model

In this formula, x is the pixel coordinate, $LE(I(\mathrm{x});\alpha)$ is the enhanced result of input $I(\mathrm{x})$, $\alpha \in [-1,1]$ is the trainable parameter of the curve, the LE curve described in (6) can be repeatedly employed to enable the capture of additional details in order to tackle the enhancement of more challenging low-light input images. Building on this idea, the subsequent high-order curve equation is derived:

$$LE_n(x) = LE_{n-1}(x) + \alpha LE_{n-1}(x)(1 - LE_{n-1}(x)) \quad (7)$$

In this formula, n is the number of iterations to control curvature. When n =1, (7) and (6) are equivalent, and higher-order curves have stronger adjustment ability compared to quadratic curves.

(2) Depth Curve Estimation Network (DCE-Net)

DCE-Net is a deep convolutional network that enhances low-light images by transforming them into a pixel-level curve parameter map of high-order curves. It has seven convolutional layers with ReLU activation between adjacent layers and a Tanh activation after the last layer. To maintain the connection between adjacent pixels, DCE-Net does not use downsampling and batch normalization layers. With only 79416 trainable parameters, DCE-Net is very lightweight.

(3) Non-Reference Loss Functions

DCE-Net uses differentiable non-reference loss functions to achieve "zero-reference" learning, which assesses and improves image enhancement based on the magnitude of these functions. The algorithm employs four types of losses, including Spatial Consistency Loss, Exposure Control Loss, Color Consistency Loss, and Illumination Smoothness Loss, to train DCE-Net.

C. Improved Defogging Framework

The improved defogging enhancement network framework is shown in Fig. 6. Firstly, the improved loss function FFA-Net is used to remove fog from the input haze image. Then, to address the issue of reduced image brightness after removing fog, and the poor performance of original FFA-Net under low lighting conditions, a low illumination enhancement module is used to enhance the low brightness image. Finally, the required clear fog free image is obtained through by output image.

IV. Experiment And Analysis

A. Dataset Selection and Model Training

To ensure high-quality image defogging results at the pixel level and for human visual perception, this paper uses the Cityscapes_foggy dataset for both training and testing, which is a synthesized fog dataset, and the real haze dataset RTTS for additional testing. The Cityscapes_foggy dataset is generated using three different atomization parameters on the Cityscapes dataset, which comprises 2957 training sets, 500 validation sets, and 1525 test sets. Based on the level of atomization, the Cityscapes_foggy dataset can be categorized into no fog (original image), thin fog, medium fog, and thick fog datasets, each containing an equal number of images. The RTTS dataset consists of 4322 real haze images captured in outdoor natural haze environments with varying concentrations.

The paper was trained using the PyTorch deep learning framework in Python. After multiple rounds of experimentation, the hyperparameters were set as follows: the value of α in (5) was set to 0.4, batch size was set to 16, and initial learning rate was set to 0.01. The Adam algorithm was used to dynamically adjust the learning rate with a momentum of 0.937. The model was trained for 200 rounds.

To increase the number of training samples and enhance the model's generalization ability, data augmentation is performed before feeding the image into the network for training. The input image is randomly cropped, rotated, and its brightness and contrast are randomly adjusted, enabling the dehazing model to handle various types of images more accurately.

B. Analysis of Experimental Results

In order to highlight the superiority of the defogging algorithm in this paper and to conveniently and intuitively observe the differences in defogging effects of different defogging algorithms, in addition to comparing the improved FFA-Net defogging algorithm with the original FFA-Net defogging algorithm, this paper also compared the Retinex algorithm, DCP algorithm, and AOD-Net algorithm, and conducted experimental comparisons on two datasets.

Citscapes_foggy is a synthetic fog dataset, where each haze image has its corresponding original clear image. To measure the defogging effect, reference evaluation indicators such as PSNR and SSIM are utilized. Higher PSNR and SSIM values indicate better image quality. To compare the performance of different models more intuitively, five images were randomly selected from the test set to display the dehazing effect. Additionally, PSNR and SSIM were calculated with the original hazy input image. Finally, the average indicators of different models in the entire test set are calculated. The defogging results of different defogging

Fig. 7. The comparison image after dehazing on Cityscapes_foggy dataset

TABLE I
COMPARISON OF PSNR AND SSIM OF DIFFERENT ALGORITHMS ON CITYSCAPES_FOGGY DATASET

Metrics	Method	(p1)	(p2)	(p3)	(p4)	(p5)	Test set average
PSNR	Input	13.934	13.387	14.197	14.227	13.687	14.058
	Retinex	14.845	14.806	14.923	14.527	14.855	14.842
	DCP	16.027	16.553	15.988	16.176	16.569	16.659
	AOD-Net	19.743	19.165	19.060	19.022	19.810	19.362
	FFA-Net	22.147	22.815	22.177	22.975	22.433	22.496
	Ours	25.040	24.112	24.074	25.155	24.491	24.522
SSIM	Input	0.616	0.623	0.624	0.613	0.616	0.620
	Retinex	0.630	0.626	0.633	0.634	0.635	0.633
	DCP	0.654	0.643	0.651	0.635	0.589	0.654
	AOD-Net	0.684	0.681	0.679	0.683	0.678	0.681
	FFA-Net	0.721	0.731	0.719	0.722	0.713	0.722
	Ours	0.749	0.751	0.748	0.746	0.752	0.750

algorithms in Citscapes_foggy are shown in Fig. 7. The statistics of PSNR and SSIM evaluation indicators for different algorithms are shown in Table I.

From Fig. 7, it can be seen that the Retinex algorithm and DCP algorithm have poor defogging effects, with a large amount of fog remaining on the image. The Retinex algorithm also leads to a decrease in image brightness, resulting in color oversaturation. Although the AOD-Net algorithm can remove fog to a certain extent, the defogging effect is not thorough, and the edge information of the image is lost. Instead, the algorithm in this paper has better visual perception of image brightness and contrast, and the defogging effect is the best. The evaluation of image quality after defogging not only takes into account human subjective visual perception, but also relies more importantly on objective evaluation indicators. From Table Ⅰ, it can be seen that the PSNR and SSIM of our method are significantly higher than other methods on different input sample plots, with significant advantages. Compared to the benchmark FFA-Net algorithm, the average values of PSNR and SSIM in our method on the test set were 2.026 and 0.028 higher, respectively, proving the effectiveness of our improved method and the superiority of defogging effect.

The algorithm proposed in this paper has demonstrated exceptional performance on synthetic fog datasets. Moving forward, the defogging effectiveness of the algorithm will be

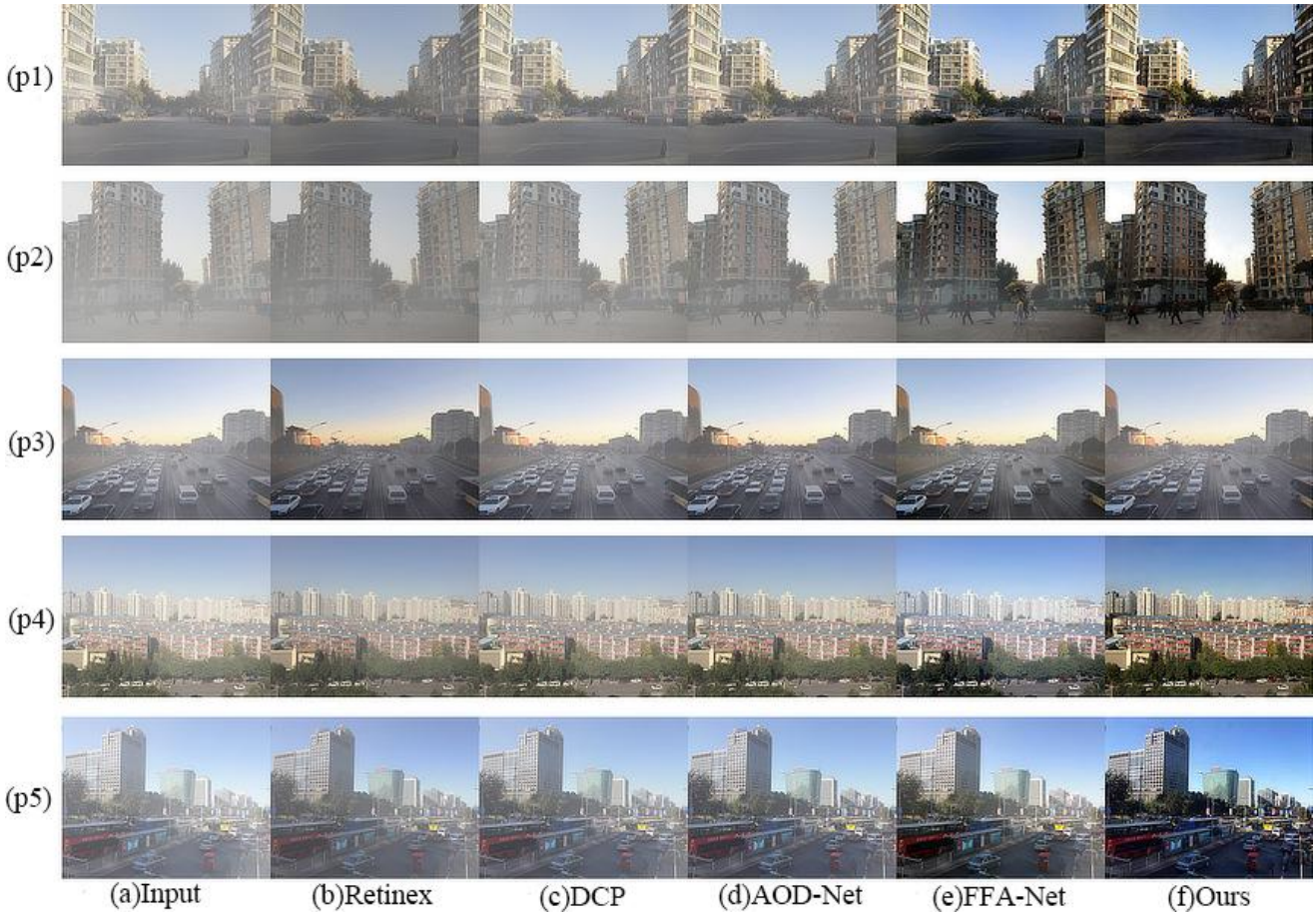


Fig. 8. The comparison image after dehazing on RTTS datase

TABLE II
COMPARISON OF PSNR AND SSIM OF DIFFERENT ALGORITHMS ON RTTS DATASET

Metrics	Method	(p1)	(p2)	(p3)	(p4)	(p5)	Test set average
NIQE	Input	18.936	18.257	17.197	19.227	18.387	18.036
	Retinex	16.734	16.896	16.323	16.648	15.898	16.858
	DCP	14.527	15.543	15.528	16.638	15.169	15.456
	AOD-Net	13.243	13.265	14.360	13.231	14.010	14.162
	FFA-Net	11.247	13.614	10.577	12.375	13.253	12.459
	Ours	10.030	9.112	9.046	11.155	9.364	9.342
PIQE	Input	22.626	24.534	19.524	23.823	21.686	22.631
	Retinex	21.630	22.626	18.539	20.621	19.835	20.641
	DCP	18.756	19.643	17.496	19.635	17.949	18.689
	AOD-Net	16.684	17.612	16.789	17.586	16.946	16.632
	FFA-Net	18.721	14.720	13.612	14.634	13.371	13.124
	Ours	11.684	11.258	10.862	9.746	8.732	10.434

evaluated on real fog datasets. Fig. 8 displays a selection of image defogging results. Similarly, considering the absence of a baseline clear image in the dataset, two non-reference indicators were chosen to assess the defogging efficacy in real foggy environments, namely Natural Image Quality namely Natural Image Quality Evaluator (NIQE) [23] and Perception based Image Quality Evaluator (PIQE) [24]. Both of these indicators suggest that higher image quality is indicated by higher values. The results of the comparison are presented in Table II.

From Fig. 8, it can be seen that our defogging algorithm is still effective in real foggy environments and is superior to other algorithms. The defogged images are also more excellent in color and contrast. Analyzing Table II reveals that our method demonstrates smaller NIQE and PIQE values compared to other methods for different input sample plots, indicating significant advantages. When compared to the benchmark FFA-Net algorithm, our method achieves a lower average NIQE of 3.117 and a lower average PIQE of 2.690 on the test set. This observation provides further evidence of the effectiveness and superiority of our enhanced FFA-Net defogging method in real foggy environments.

C. Ablation Experiment

This paper conducted ablation experiments to confirm the significance of each module in the proposed defogging model and to further validate the effectiveness of the method. The original FFA-Net, FFA-Net with Zero-DCE added, FFA-Net with improved loss function, and FFA-Net with Zero-DCE added and improved loss function (ours) were compared experimentally on the dataset Citscape_foggy and dataset RTTS, respectively. In addition to the defogging effect evaluation indicators mentioned above, in order to test the impact of newly added modules on model inference speed, we also compared the FPS of different methods. The comparison results are shown in Table III and Table IV,

respectively.

TABLE III
COMPARISON OF ABLATION EXPERIMENTS ON CITYSCAPES FOGGY

Metrics	FFA-Net	FFA-Net+ Zero-DCE	FFA-Net+ improved loss	ours
PSNR	22.491	23.512	23.623	24.513
SSIM	0.720	0.736	0.739	0.751
FPS	136	128	135	128

TABLE IV
COMPARISON OF ABLATION EXPERIMENTS ON RTTS

Metrics	FFA-Net	FFA-Net+ Zero-DCE	FFA-Net+ improved loss	ours
NIQE	12.449	11.121	10.979	9.251
PIQE	13.119	11.894	11.793	10.422
FPS	137	131	137	130

From Tables III and IV, it can be seen that after adding Zero-DCE and improving the loss function, the indicators NIQE and PIQE are significantly higher than before improvement. After improving the loss function, there is almost no change in FPS, it shows that modifying the loss function will only affect the training speed but not the inference speed. After adding Zero-DCE, FPS decreased slightly. In summary, our model significantly improves the dehazing effect while slightly reducing the inference speed, indicating that our model can adapt to different types of foggy datasets, and proving the effectiveness and necessity of each improved module of our model.

V. CONCLUSION

In this paper, we propose an improved FFA-Net defogging algorithm, introduce the structural characteristics and working principle of FFA-Net, and improve and optimize it to address its shortcomings. In order to enable the network to learn the best pixel and visual features, this paper combines the L1 loss and SSIM loss of FFA-Net, and enhances the defogging effect under low brightness conditions, This paper introduces the low illumination enhancement module Zero-DCE to enhance the image after defogging. By conducting comparative experiments, it was discovered that the proposed method in this paper produces more realistic and natural results compared to other methods. This is supported by the ideal objective indicators and visually pleasing outcomes. Therefore, it can be concluded that the algorithm proposed in this paper outperforms other classic defogging algorithms.

REFERENCES

[1] Sharma N, Kumar V, and Singla S K, "Single image defogging using deep learning techniques: past, present and future," *Archives of Computational Methods in Engineering*, vol. 28, pp. 4449-4469, 2021.

[2] Li C, Guo C, Han L, et al, "Low-light image and video enhancement using deep learning: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligenc*, vol. 44, no.12, pp. 9396-9416, 2021.

[3] Zhang, Junkai, et al, "Single-Image Defogging Algorithm Based on Improved Cycle-Consistent Adversarial Network," *Electronics*, vol. 12, no.10, pp. 2186-2210, 2013.

[4] Yuan B, Yang Y, and Zhang B, "Single Image Defogging Method based on Deep Learning," 2017 International Conference on Mechanical, Electronic, Control and Automation Engineering (MECAE 2017), 2017, Atlantis Press, pp. 126-131.

[5] Dhal K G, Das A, Ray S, et al, "Histogram equalization variants as optimization problems: a review," *Archives of Computational Methods in Engineering*, vol. 28, pp. 1471-1496, 2021.

[6] Liu, Shouxin, et al, "Retinex-based fast algorithm for low-light image enhancement," *Entropy*, vol. 23, no.6, pp. 746-760, 2021.

[7] Singh G, Mittal A, "Various image enhancement techniques-a critical review," *International Journal of Innovation and Scientific Research*, vol. 10, no.2, pp. 267-274, 2014.

[8] Land E H, and McCann J J. "Lightness and retinex theory," *Josa*, vol. 61, no.1, pp. 1-11, 1971.

[9] Bansal B, Sidhu J S, and Jyoti K, "A Review of Image Restoration based Image Defogging Algorithms," *International Journal of Image, Graphics and Signal Processing*, vol. 9, no.11, pp. 62-74, 2017.

[10] Xinfu, Chen, Jianlong, et al, "Image Defogging Approach Based on Dark Channel Prior with Feedback Mechanism," *Journal of Information and Computational Science*, vol.11, no.16, pp. 5855-5862, 2014.

[11] McCartney E J, "Optics of the Atmosphere: Scattering by Molecules and Particles," *Physics Today*, vol. 30, no.5, pp. 76-77, 1977.

[12] Narasimhan S G, and Nayar S K, "Interactive (de) weathering of an image using physical model," IEEE Workshop on Color and Photometric Methods in Computer Vision, 2003, France, vol. 6, no.6.4, pp. 1-8.

[13] Tan K K, and Oakley J P, "Physics-based approach to color image enhancement in poor visibility conditions," *JOSA A*, vol. 18, no.10, pp. 2460-2467, 2001.

[14] He K, Sun J, and Tang X, "Single image haze removal using dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no.12, pp. 2341-2353, 2010.

[15] Sun Y B, Xiao L, and Wei Z H, "Method of defogging image of outdoor scenes based on PDE," *Journal of System Simulation*, vol. 19, no.16, pp. 3739-3769, 2007.

[16] Cai B, Xu X, Jia K, et al, "Dehazenet: An end-to-end system for single image haze removal," *IEEE Transactions on Image Processing*, vol. 25, no.11, pp. 5187-5198, 2016.

[17] Li B, Peng X, Wang Z, et al, "Aod-net: All-in-one dehazing network" Proceedings of the IEEE International Conference on Computer Vision, 2017, Venice, France, pp. 4770-4778.

[18] H Zhang, and V M Patel, "Densely connected pyramid dehazing network," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, Salt Lake City, USA, pp. 3194-3203.

[19] Qin X, Wang Z, Bai Y, et al, "FFA-Net: Feature fusion attention network for single image dehazing," Proceedings of the AAAI Conference on Artificial Intelligence, New York, USA, 2020, vol. 34, no.07, pp. 11908-11915.

[20] Bang-guWU, Su-linZHANG, HongSHI, et al, "Multi-Branch Structure Based Local Channel Attention with Uncertainty," *Acta Electronica Sinica*, vol. 50, no.02, pp. 374-382, 2022.

[21] Jia F, Ma L, Yang Y, et al, "Pixel-attention CNN with color correlation loss for color image denoising," *IEEE Signal Processing Letters*, vol. 28, pp. 1600-1604, 2021.

[22] Guo C, Li C, Guo J, et al, "Zero-reference deep curve estimation for low-light image enhancement," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, Seattle, USA, pp. 1780-1789.

[23] Wu L, Zhang X, Chen H, et al, "VP-NIQE: An opinion-unaware visual perception natural image quality evaluator," *Neurocomputing*, vol. 463, pp. 17-28, 2021.

[24] Chen S D, "Human visual perception-based image quality analyzer for assessment of contrast enhancement methods," *Int. Arab J. Inf. Technol*, vol. 13, no.2, pp. 238-245, 2016.