

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Iterative Residual Network for Image Dehazing

ZHEN HUA^{1,2}, GUODONG FAN^{1,2}, AND JINJIANG LI.^{1,2},

¹School of Computer Science and Technology, Shandong Technology and Business University, Yantai 264005

²Co-innovation Center of Shandong Colleges and Universities: Future Intelligent Computing, Yantai 264005

Corresponding author: Zhen Hua (e-mail: huazhen66@foxmail.com).

ABSTRACT In this paper, we propose an Iterative Residual Network. By designing the calculation unit, we put the hazy image into the calculation unit to perform an iterative operation that can stitch the hazy image in stages with the unit output and substitute it into the calculation. After multiple iterations, a clean image can be generated. We introduce Long Short-Term Memory network and Residual ideas in the design process of the computing unit to further optimize the model. The Long Short-Term Memory network can be used to connect computing units at different stages. The use of residual block connection in the deep processing of the computing unit can preserve the original features of the image and prevent the model from overfitting. This model directly generates hazy-free images in an end-to-end manner and does not rely on the estimations of the transmission map and atmospheric light. Experiments show that Iterative Residual Network can effectively remove the haze in the image. In the test of the synthetic dataset and the real dataset, Iterative Residual Network is superior to the existing methods in terms of PSNR, SSIM, FADE and subjective visual effects.

INDEX TERMS Iteration, Residual, Long Short-Term Memory, Image Dehazing

I. INTRODUCTION

AS light is scattered and absorbed by small particles in the atmosphere, the contrast and saturation of images in haze weather will be seriously affected which will lead to color deviation and make it difficult for outdoor vision system to work normally. Therefore, how to make hazy image clearer has become a hot topic for researchers and also a difficult problem in Computer Vision and Image Processing.

Most of the previous studies improve the display effect of hazy image by improving the image contrast or reducing the pattern noise. Among them, representative methods include Histogram Equalization, Wavelet Transform and Homomorphic Filtering, etc. However, because these methods do not take the imaging mechanism into consideration, the processed image often loses some of the original image details. In addition, the difficulty in parameter adjustment also increases the complexity of these algorithms.

Existing research usually uses the following models to model the imaging process of hazy images:

$$I(x) = T(x)J(x) + A(1 - T(x)) \quad (1)$$

$$T(x) = e^{-\beta d(x)} \quad (2)$$

where $I(x)$ is a hazy image; $J(x)$ is a clean image; A is the global atmospheric light; $T(x)$ is the image transmission rate; x is the position of the pixel in the image; β is the atmospheric scattering coefficient and $d(x)$ is the scene depth.

It can be seen from formula (1) and (2) that if we can get a suitable transmission map and global atmospheric light, the hazy image can be restored to a haze-free image. Based on the atmospheric scattering model [1] [2] [3], early research mainly used apriori methods to estimate the atmospheric light and transmission map, such as the Dark Channel Prior algorithm by HE et al. [4]. The apriori algorithm can get good results in some cases, but when the real situation is contrary to the estimated situation, the inaccurate estimation of the transmission map will result in unsatisfactory dehazing effect. With the application and development of deep learning theory [5] [6] in the field of computer vision, some scholars have begun to use convolutional neural networks to estimate transmission maps and atmospheric light. For example, Cai et al. [7] proposed DehazeNet, which is dedicated to learning the relationship between hazy images and image transmission rates. Li et al. [8] AOD-NET converts the transmission map and atmospheric light into a unified variable K, and uses a modified atmospheric physical model directly

generates hazy-free images through CNN. The application of deep convolutional neural networks has greatly promoted the development of image dehazing technology. However, an important existing problem is that the neural network method based on the physical model is highly dependent on the transmittance and atmospheric light, which means that when the model parameter estimation is not accurate, the output result will be greatly affected.

Recently, some scholars regard the image dehazing process as a conversion problem between images, rather than relying on the estimation results of atmospheric physical models. For example, Qu et al. [9] drew on the EPDN structure proposed by Generative Adversarial Networks. Liu et al. [10] used the attention mechanism to propose GridDehazeNet which promotes the development of multi-scale estimation methods. Dong et al. [11] proposed a Multi-Scale Boosted Dehazing Network with dense feature fusion (MSBDN) based on U-Net [12] with dense fusion of multi-scale features. Although the above methods avoid the atmospheric physical model, most of them use a general network architecture and do not further optimize the network architecture for the image dehazing problem. Therefore, the output images generally have the phenomenon of blurred details and loss of details.

In the research of image restoration technology such as image dehazing [13], rain removal [14] and image deblurring [15], it is very important to ensure that the restored image still retains detailed information. In the research field of image restoration, although the pooling layer can reduce the dimensionality of the data and perform network calculations, since simply increasing the number of layers of the network will bring the risk of overfitting and cannot retain the image detail information, making the application of the pooling layer has not received widespread attention.

This paper proposed an Iterative Residual Network. The method proposed in this paper mainly includes Dehaze Unit and model architectures for iterative calculations. The Dehaze Unit itself can be regarded as a complete end-to-end model, set the number of iterations in the model architecture so that the hazy image is iterated in the Dehaze Unit, and the output of the Dehaze Unit after each iteration is spliced with the network input as the input for the next iteration. Iterative operation does not increase the number of network parameters in multiple iterations. Through experiments, it is found that different iteration times have different dehaze capabilities. In the design of the Dehaze Unit, the multi-layer convolution is used to extract the deep-level features. We use the residual block connection method to improve the fitting ability of the model and ensure that the model can be reliably trained.

The method we proposed does not rely on atmospheric physical models, so it avoids the problem of poor dehazing effect caused by inaccurate estimates of transmittance and atmospheric light. We use paired images for training and build channels between hazy and hazy-free images. Experiments show that the results of this method are better than existing methods in Peak signal-to-noise Ratio (PSNR), Structural

SIMilarity (SSIM) [16] and hazy density estimation model (FADE) [17]. The main contributions of this article may be the following points:

1. Designed Iterative Residual Network (IRN) that can directly generate hazy-free images. This method combines iteration and residuals for image dehazing. The restored image has greater advantages in terms of visual perception and image detail retention.

2. According to experiments, the higher the number of iterations of the IRN structure, the better the dehazing effect. The residual structure introduced in this method can effectively prevent the model degradation problem caused by the deep network structure.

3. This method applies the skip-connection idea [18] to the iterative process of IRN. Since the output at different moments can be combined, the fitting effect of the model is better. As we expected, the dehazing effect of this method is better than the pure iterative model in SSIM, PSNR, FADE and visual perception.

II. RELATED WORK

The research on image dehazing technology first appeared in 1992, Bissonnette [19] described a new and efficient method of calculation of the point spread and modulation transfer functions (PSF and MTF) caused by aerosol forward scattering. After that, Oakley [20] proposed a temporal filter structure to enhance the signal-to-noise ratio in the image. Cantor [1] studied the light scattering of atmospheric molecules, aerosols and cloud particles. Subsequently, Nayar and Narasimhan tried to mathematically describe the phenomenon of atmospheric scattering and finally put forward the atmospheric physical model, which laid the foundation for the research in the field of image dehazing [2] [3] [21].

Image dehazing is an ill-posed problem. In this section, we will introduce the mainstream methods of image dehazing in detail. The existing image dehazing methods can be roughly divided into image dehazing methods based on prior algorithms and image dehazing methods based on learning.

A. IMAGE DEHAZING METHODS BASED ON PRIOR ALGORITHMS

The representative scholar in this research field is HE et al. [4] The dark channel prior method (DCP) he proposed for estimating the transmission image believes that the local minimum of the dark channel of a hazyless image is close to zero. This method is groundbreaking, but its high computational complexity leads to extremely low dehazing efficiency. Schechner [22] et al. first proposed a method of using two pictures with different polarization directions to estimate the degree of polarization, ambient light intensity, and transmittance distribution of the image, and to achieve dehazing. Based on the atmospheric scattering model, Fattal's et al [23] method assumed that the shadow of the object surface and the transmittance of the ring are statistically locally uncorrelated and then the original image color is obtained through the Markov random field, but Markov Random Field also has

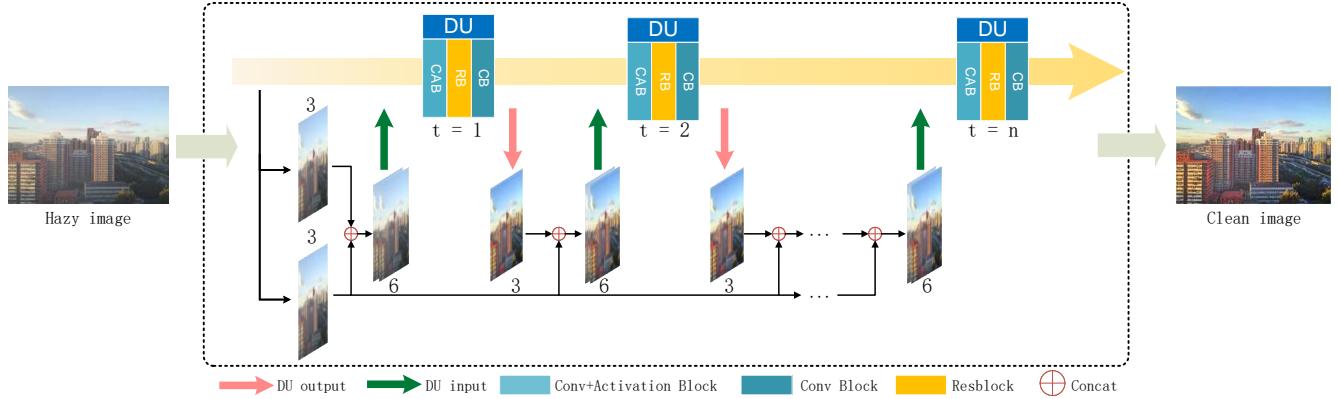


FIGURE 1: The architectures of Iterative Residual Network, inputs foggy images into the network, and the network directly outputs clean images. The model uses the iterative process of the dehazing unit. In the iterative process, the output of the previous iteration and the network input are connected in series as the input of the next iteration. The structure of the Dehaze Unit is shown in Fig.2.

the problem of high computational complexity. Tan et al [24] assumed that the ambient light in the local area of the image is constant and enhances the contrast of the image, and used the image segmentation theory to estimate Optimal lighting. Although this method can restore the structure and details of the image, the restored image is often too saturated and halo occurs on the boundary of the deep image. Berman et al [25] believes that the color of a hazy-free image can be fitted with hundreds of different colors, and proposes a non-local prior method to describe the characteristics of a clear image. In short, the image dehazing method based on apriori algorithm cannot be applied to all scenes. Once the estimation result does not match the actual situation, the image dehazing effect will deteriorate. Shen et al . [26] provided depth chromaticity compensation regularization for transmission images and chromaticity depth compensation regularization for image dehazing, the proposed iterative image dehazing method with polarization used these two joint regularization schemes and transmission the relationship between the image and the dehazing image.

B. IMAGE DEHAZING METHODS BASED ON LEARNING.

In recent years, deep learning theory has developed well in the field of image processing, especially in the fields of semantic segmentation [27], object detection [28], and image super-resolution [29]. Some scholars have begun to study how to use Deep Convolutional Neural Network (DCN) [5] for image dehazing. The existing learning-based methods can be divided into two categories: one is to estimate the transmission map and atmospheric light through DCN [7] [30] [31] [8] [32] ; the other is to directly output the hazy-free image end-to-end [33] [10] [34] [9] [35] [34] [36] [37] [11] .

Cai et al. [7] proposed the DehazeNet. This method mainly used convolutional neural network to extract hazy-related features in the image (such as dark primary color features,

color attenuation features, maximum contrast features, etc.) to optimize the model's estimate of transmittance. In addition, this method also presets atmospheric light data based on experience. Ren et al. [30] designed a set of coarse-scale and fine-scale networks to predict the image transmission rate (MSCNN) separately and use multi-scale fusion to complete image dehazing. Li et al. [8] proposed an AOD-Net that simplified the unknown in the atmospheric scattering model to a coefficient K and learned the relationship between hazy images and K through CNN. Bharath et al. [38] mainly used the residual block to estimate the image transmission rate and the global atmospheric light for the atmospheric scattering model to recover the hazy image, and they also use a single-scale discriminator for generative adversarial learning (DehazeGAN). When the transmission map and the atmospheric light estimate are not accurate, the image dehazing will be greatly affected. The Gated Fusion Network (GFN) [39] proposed by Ren et al. used white balance, contrast enhancement and gamma correction to preprocess hazy images, and used CNN to learn the confidence maps corresponding to three preprocessed images, and then through multi-scale fusion Obtain a hazy-free image. Qu et al. [9] proposed an enhanced Pix2Pix dehazing network (EPDN) which used a phased dehazing module to enhance the dehazing effect. Chen et al [34]. proposed an end-to-end gated context aggregation network (GCANet), which used a smooth expansion technique to eliminate gridding artifacts caused by negligible parameters of the expanded convolution kernel, and leverage a gated sub- network to fuse the features from different levels. Liu et al. [10] proposed GridDehazeNet by integrating the attention mechanism into multi-scale estimation. Qin et al. [37] proposed an end-to-end feature fusion attention network (FFA-Net), through an attention-based feature fusion structure that can retain shallow information and transfer it to the deep. Dong et al. [11] proposed a Multi-Scale Boosted Dehazing Network with dense feature fusion based on the U-Net

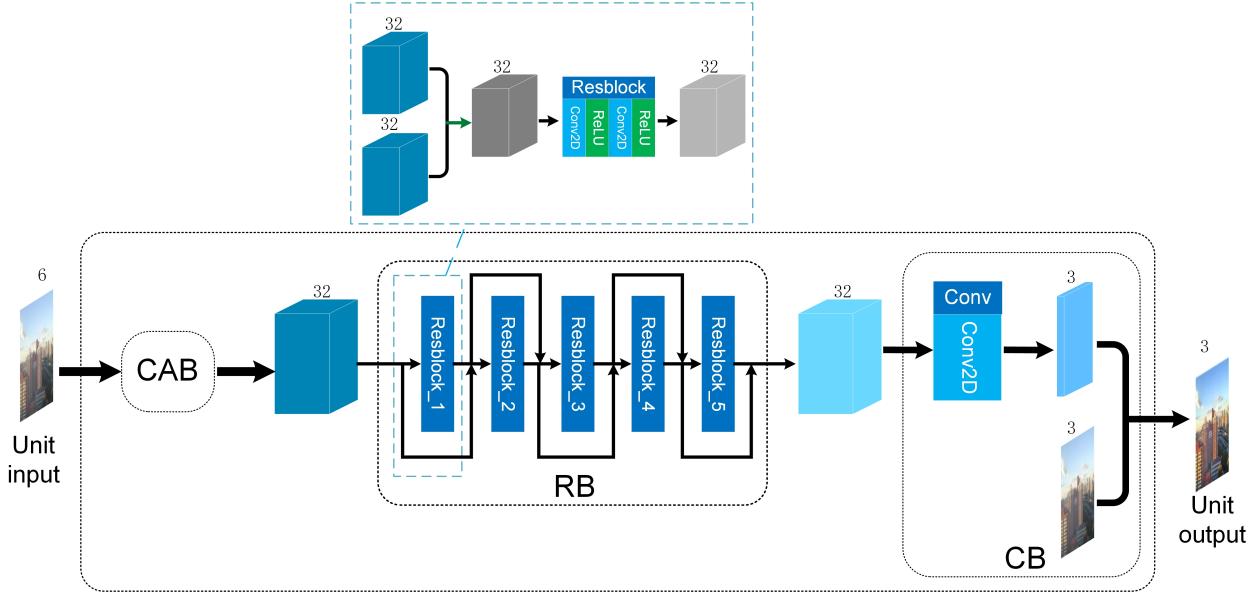


FIGURE 2: The internal structure of Dehaze Unit consists of Convolution and Activate Block (CAB), Residual Block (RB) and Convolution Block (CB). The specific structure of CAB is shown in Fig.3.

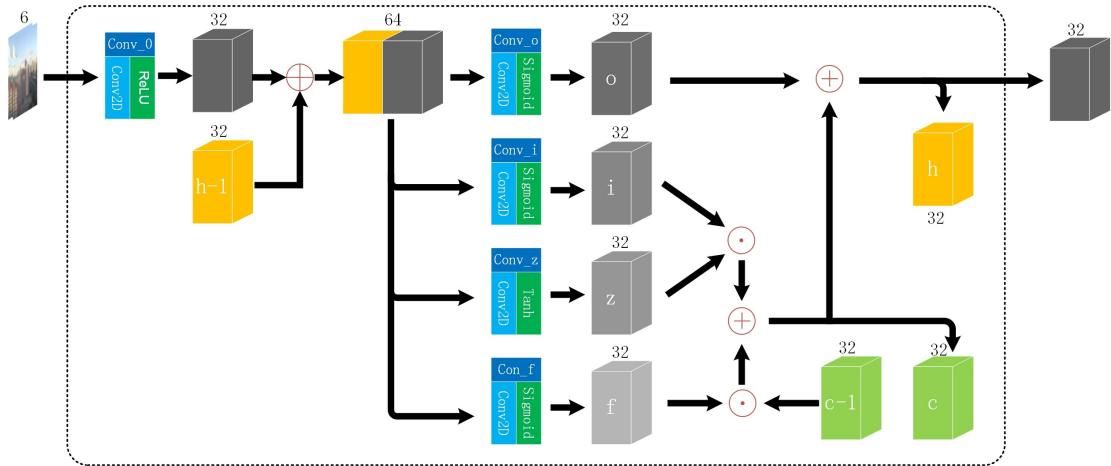


FIGURE 3: Internal structure diagram of Convolution and Activate Block (CAB). CAB is the first calculation part of DU, which is responsible for extracting the features of DU input. The input of CAB is a hazy image, and the output is a 32-channel feature map.

[12] architecture. This method is designed based on the two principles of boosting and error feedback, which shows that they are suitable for dehazing. By adding the "Strengthen-Operate-Subtract" enhancement strategy to the decoder of the proposed model, a simple and effective enhancement decoder is developed to gradually restore the hazy-free image.

III. MODEL CONSTRUCTION

In this section, we will introduce our proposed Iterative Residual Network (IRN) in detail through three aspects: network structure, model input and output, and loss function. As shown in Fig.1, the input network is a hazy image. After the image is concatenated, it is input to the Dehaze Unit (DU)

, and the result of the previous step is concatenated with the DU and then enters the DU unit again. Iterates n times and finally outputs Clean image.

A. ITERATIVE NETWORK.

From previous studies, it can be found that simply increasing the depth of the network can not guarantee a better dehazing effect, and there are too many parameters that are difficult to train and the risk of overfitting. Kim et al [40] proposed the Deep-Recursive Convolutional Network for image super-resolution processing . The recursive method in DRCN provides us with new ideas, so when we design the deep network model structure, we adopt in an iterative way, an

TABLE 1: THE ARCHITECTURES OF THE IRN

Formulation		Type	Id	Input Size	Num	Size	Strided	Output Size
Input		Concat	Concat0	(3,640,480) (3,640,480)	—	—	—	(6,640,480)
DU	CAB	Conv+ ReLU	Conv0	(6,640,480)	32	3x3	1	(32,640,480)
		Concat	Concat1	(32,640,480) (32,640,480)	—	—	—	(64,640,480)
		LSTM	Conv+Sigmoid	Conv_o	(32,640,480)	32	3x3	1 (32,640,480)
			Conv+Sigmoid	Conv_i	(32,640,480)	32	3x3	1 (32,640,480)
			Conv+Tanh	Conv_z	(32,640,480)	32	3x3	1 (32,640,480)
			Conv+Sigmoid	Conv_f	(64,640,480)	32	3x3	1 (32,640,480)
			Comput	— (32,640,480) (32,640,480) (32,640,480) (32,640,480)	—	—	—	(32,640,480)
		ResBlock	Comput	— (32,640,480) (32,640,480)	—	—	—	(32,640,480)
			Conv+ReLU	— (32,640,480)	32	3x3	1	(32,640,480)
			Conv+ReLU	— (32,640,480)	32	3x3	1	(32,640,480)
	RB	ResBlock
		ResBlock	Comput	— (32,640,480) (32,640,480)	—	—	—	(32,640,480)
			Conv+ReLU	— (32,640,480)	32	3x3	1	(32,640,480)
			Conv+ReLU	— (32,640,480)	32	3x3	1	(32,640,480)
		CB	Conv	Conv1	(32,640,480)	3	3x3	1 (3,640,480)
			Comput	— (3,640,480) (3,640,480)	—	—	—	(3,640,480)
Output		—						(3,640,480)

overall computing unit is designed so that the features are iterated within this computing unit. The parameters of this overall computing unit are shared during different iterations, thereby significantly reducing network model parameters and speeding up model training.

B. IRN DESIGN.

The main idea of the Iterative Residual Network (IRN) (see Fig.1) that we first proposed is to input the Hazy image into the iterable Dehaze Unit (DU) we designed to process the results (see Fig.2), the result of the DU will be processed in the DU again, and iterate until the clean image is obtained. Different iterations will have different computational complexity and effects, and the number of iterations should be limited. Therefore, we need to define the number of iterations in advance. Obviously, the higher the number of iterations, the higher the computational complexity of the network and the difficulty of training. With the increase, after the input data is repeatedly convolved and activated, the original details of the image will inevitably be lost, and then the training results will be over-fitted and difficult to generalize. The input image of the model will jump to the DU the output of DU enters the next iteration after concatenation. Experiments have proved that this method can indeed retain the original details of the image, so that the network with a higher number of iterations can be effectively trained. The main calculation process of DU as

$$T_{t+1} = CB\{RB[CAB(T_t \oplus T_{t-1})]\} \quad (3)$$

where T represents the output value of the DU after a certain iteration. When $t=0$, it represents the input image. At this time, the input image is copied and processed after Concat. CAB, RB and CB are the three calculation parts in DU.

Convolution and Activate Block(CAB). CAB is the first calculation part of DU, which is responsible for extracting the features of DU input. The first step of CAB is to use 32 convolution kernels with a size of 3x3 to extract features, and perform non-linear activation through the activation function. The output channel is 32 feature maps. We consider that under high iteration times, multiple convolution operations of the model may weaken some of the original features of the image. Therefore, we introduce LSTM [41] here to connect the CAB at different iterations. We consider that under high iteration times, multiple convolution operations of the model may weaken some of the original features of the image. Therefore, we connect the LSTM layer here to connect the CAB at different iterations to connect the previous iteration. The feature map output by CAB ($h-1$, the transmission of the previous step) is introduced here for splicing (in the first iteration, h is the initialized 32-channel feature map), and the spliced feature maps are respectively rolled four times Product, four feature maps of o , i , z and f are obtained. i , z , f and c_{t-1} are calculated according to formula (4), and the states c_t , c_t and o are obtained according to formula (5), ht is the current iteration time CAB output.

$$c_t = f \times c_{t-1} + i \times z \quad (4)$$

$$h_t = o \times \text{Tanh}(c_t) \quad (5)$$

Residual Block(RB). RB is the second part of DU, which receives the result of CAB and further extracts deep features. For this reason, the structure of our residual connection [18] uses five residual blocks to connect, and all residual block structures input and output 32-channel feature maps. The residual block contains two pairs of "convolution kernel + ReLU" combination, the size of the convolution kernel is 3×3 , the padding is 1, the output of the residual block is a 32-channel feature map, and the input of the residual block It is connected with the output jump and added as the input of the next residual block. The calculation process is shown in Fig.2, and the calculation method is shown as

$$f_{n+1}(x) = \text{ReLU}(f_n(x) + f_{n-1}(x)) \quad (6)$$

where f_n is the output of the current residual block, f_{n-1} is the input of the current residual block or the output of the previous residual block, f_{n+1} is the input of the next residual block, and the previous residual block and the current residual block the difference blocks are added and a nonlinear activation operation is performed by the ReLU activation function.

Convolution Block(CB). CB is the final calculation part of the DU, responsible for final processing of the feature results extracted by the RB of the DU and outputting the image as the result of the DU. The specific process is to receive the feature map output by the RB, convolve it into a channel 3 feature map with a size 3×3 convolution kernel, and add it to the model input image (hazy image) to obtain the final image.

C. LOSS FUNCTION.

In recent studies, the loss function of the dehazing network emerges endlessly, most of which are in mixed form, For example, MSE+SSIM [42], and the weighted summation of the adversarial loss [43], the feature matching loss [44], the perceptual loss [45] and the fidelity loss [46] , the loss index of the image is integrated into the loss function, which can help the model achieve Good results, but greatly increase the risk of hyperparameter tuning. The IRN proposed in this paper benefits from the iterative idea. After many experiments, it is found that only using a single MSE Loss and negative SSIM [47] as the loss function of IRN can have an obvious dehazing effect.

IRN is an iterative network, there are t moments, that is, the output of DU after each iteration is $Y^{(1)}, Y^{(2)} \dots Y^{(t)}$, t is the number of iterations set by IRN, and $Y^{(t)}$ is the model's the final output, so the MSE loss of IRN can be expressed as

$$L_t = \frac{1}{2M} \sum_{m=1}^M (Y_m - Y_m^{(t)})^2$$

Its egative SSIM loss is

$$L_t = -\frac{1}{2M} \sum_{m=1}^M \frac{(2\mu_{Y_m}\mu_{Y_m^{(t)}} + \theta_1)(2\sigma_{Y_m Y_m^{(t)}} + \theta_2)}{(\mu_{Y_m}^2 + \mu_{Y_m^{(t)}}^2 + \theta_1)(\sigma_{Y_m}^2 + \sigma_{Y_m^{(t)}}^2 + \theta_2)}$$

where M is the total number of samples, m is the current sample, μ_{Y_m} is the average of Y_m , $\mu_{Y_m^{(t)}}$ is the average of $Y_m^{(t)}$, σ_{Y_m} is the variance of Y_m , $\sigma_{Y_m^{(t)}}$ is the variance of $Y_m^{(t)}$, $\sigma_{Y_m Y_m^{(t)}}$ is the covariance of Y_m and $Y_m^{(t)}$, θ_0 and θ_1 are constants, θ_0 and θ_1 are to avoid system instability caused by the denominator being 0. The value range of -SSIM is [-1,0]. The idea of using -SSIM as the loss function is that when the structural similarity between the input and the output is greater, the value of SSIM is larger, and the opposite number represents the smaller the loss.

IV. EXPERIMENTAL RESULTS

In this section, we will verify the effectiveness of our proposed method on synthetic data sets and pictures of real scenes. We conducted an ablation study and verified the influence of the number of iterations and loss function on the model by comparing the peak signal-to-noise ratio (PSNR), structural similarity (SSIM) and hazy density estimation model (FADE). In addition, the proposed method Compared with the current advanced methods, mainly including D-CP, Hazy-lines, AOD-NET, GCANet, GridDehaze, FFA-Net, MSBDN.

A. EXPERIMENT SETTINGS



FIGURE 4: The ITS data set in RESIDE is defective, and the red mark box is the defective part.

Dataset. Our data set includes training set and test set. The training set contains indoor synthetic images and outdoor synthetic images. The test set includes indoor synthetic images, outdoor synthetic images and real scene images. RESIDE [48] contains images that include indoor, outdoor and real scenes. Because of RESIDE indoor synthetic hazy images, many images have image defects such as uneven haze. RESIDE contains images that include indoor, outdoor and real scenes. Since many of the indoor synthetic hazy images of RESIDE have image defects such as uneven haze (as shown in Fig.4), we choose the NYU2 dataset [49]. NYU2 contains 27257 We select 20000 images as the indoor synthetic hazy image dataset for training, and 1000 images for testing. For outdoor synthetic hazy, we use the Synthetic Outdoor Hazy Images (SOTS) dataset in RESIDE. SOTS contains 313950 images. We selected 20,000 images in the SOTS dataset and the outdoor synthetic hazy dataset, and

selected 1,000 for testing. We use 500 images in the Real-world Task-driven Testing Set (RTTS) and 100 images of different types of real scenes we collected. The real scene images are used to test the dehazing ability of the model in real scenes.

Training Detail. During training, the indoor synthetic image dataset and outdoor synthetic image dataset are used to train the model separately. The learning rate is set to 0.0001, batchsize is set to 1, epoch is set to 10, and Adam [37] is used as the optimizer. Use Pytorch as the framework and NVIDIA TITAN RTX for training.

Quality Measures. To evaluate the performance of our proposed method, we use peak signal-to-noise ratio (PSNR, see formula (9)), structural similarity (SSIM, see formula (10)) and hazy perception density evaluator (FADE) as our objective evaluation indicators.

PSNR is a full reference evaluation method, which requires reference to clean images and images to be evaluated.

$$PSNR = 10\log_{10} \frac{255}{\sqrt{|X_{in} - X_{out}|^2}} \quad (7)$$

Where X_{in} is a hazy image output X_{out} hazy model image, the larger the PSNR value, indicating the lower degree of distortion of the image, the higher the image quality, better dehazing.

SSIM is also a full-reference evaluation method, which measures the similarity between the clean image and the image to be evaluated in terms of brightness, contrast, and structure.

$$SSIM = \frac{(2\mu_{X_{in}}\mu_{X_{out}} + \theta_1)(2\sigma_{X_{in}X_{out}} + \theta_2)}{(\mu_{X_{in}}^2 + \mu_{X_{out}}^2 + \theta_1)(\sigma_{X_{in}}^2 + \sigma_{X_{out}}^2 + \theta_2)}$$

Where $\mu_{X_{in}}$ is the average of X_{in} , $\mu_{X_{out}}$ is the average of X_{out} , $\sigma_{X_{in}X_{out}}$ is the covariance of X_{in} and X_{out} , $\sigma_{X_{in}}^2$ is the variance of X_{in} , and $\sigma_{X_{out}}^2$ is the variance of X_{out} .

FADE is a hazy concentration estimation model proposed by Choi et al. For specific principles, please refer to literature [17]. This model can predict the visibility of a hazy scene from a single image, without referring to the corresponding hazy-free image, and without relying on salient objects in the scene, so we use FADE as an evaluation index to evaluate the dehazing effect.

B. ABLATION STUDY

In order to make the model achieve the best results, we use different loss functions and different iteration times to test separately with IRN to find the most suitable network configuration for IRN.

Loss Function. The choice of loss function is directly related to the effectiveness and robustness of the model. In this section, we will discuss the performance of IRN under different loss functions. We mainly used a single MSE loss and negative SSIM(-SSIM) loss for comparison tests. By using MSE and -SSIM for training in a model with a learning rate of 0.0001, epoch set to 10, and iterations from 1 to 6,

and comparing the models with the test set, it is found that a single MSE and -SSIM are used as the loss function, the purpose of image dehazing can be achieved under a high number of iterations, and the performance of PSNR, -SSIM and FADE can reach a higher level.

In the test of the model through the indoor synthetic foggy image, it can be seen by observing Fig.6 that as the number of iterations increases, the models using MSE loss and -SSIM loss are able to remove the haze and restore image details. However, under low iteration times, the image texture of the model using MSE loss is patchy, and the dehazing effect of the area with more complex image structure is poor, but the model using -SSIM loss is better than using the texture restoration and structure complex area processing MSE loss model.

In the test of the model by synthesizing the outdoor hazy image, as shown in Fig.7 and TABLE 3, the model using MSE loss has serious color shift in IRN(1) and IRN(2), and a large area will appear locally. Dark spots, IRN (3) has greatly improved the image color processing compared to IRN(2), but there are still dark spots in the image. IRN (6) has better color recovery than images with low iterations, the image details are clearer. Using the -SSIM loss model, IRN(1) has better color recovery capabilities than using MSE loss. As the number of iterations increases, the details of the image become more prominent.

Comprehensive objective evaluation and subjective visual perception, IRN using -SSIM loss is better, so follow-up tests will use -SSIM loss as the loss function.

Number of iterations. The core of the method proposed in this article is iteration. According to our assumptions, the effect of the model varies with the number of iterations. Therefore, we trained the models with iteration number of 1-10 respectively. Through objective testing of models with different iteration times, we found that as the number of iterations increases, the PSNR and SSIM of the indoor composite image and outdoor composite image are on the rise. The hazy residue degree is calculated by FADE and it is found that as the number of iterations increases As it increases, the value of FADE shows a downward trend.

In the test using the indoor composite image, from Fig.10 and Table 4, as the model increases with the number of iterations, the corresponding PSNR, SSIM and FADE have a trend of improvement. It can be seen from Fig.8 that a model with a high number of iterations is better than a model with a low number of iterations in subjective visual effects.

It can be seen from Fig.11 and TABLE 5, in the test using outdoor composite images, similar to using indoor composite images, the model with a high number of iterations is generally better than the model with a low number of iterations, and the subjective visual effect is better.

Regarding the optimal number of iterations. We make a judgment based on the objective indicators in the experimental results. It can be seen from Fig.8 and Fig.9, after the number of iterations exceeds 6, the indicators tend to be stable, and the PSNR with the number of iterations of

TABLE 2: To compare the difference between MSE loss and -SSIM loss of the model under different iteration times, the model uses indoor synthetic images for training and testing. The value in the table is the average of all images.

	IRN(1)		IRN(2)		IRN(3)		IRN(4)		IRN(5)		IRN(6)	
	MSE	-SSIM	MSE	-SSIM	MSE	-SSIM	MSE	-SSIM	MSE	-SSIM	MSE	-SSIM
PSNR	28.4209	28.4504	28.1474	28.553	28.2897	28.6705	28.3498	29.9084	29.1342	30.1463	29.6521	30.2644
SSIM	0.9042	0.9531	0.9171	0.965	0.9202	0.9701	0.9461	0.9764	0.9543	0.981	0.9651	0.9845
FADE	1.3948	1.1559	1.1203	1.0082	1.0483	0.8454	0.7632	0.6954	0.7532	0.6536	0.6354	0.5528

TABLE 3: To compare the difference between MSE loss and -SSIM loss of the model under different iteration times, the model uses outdoor synthetic images for training and testing. The value in the table is the average of all images.

	IRN(1)		IRN(2)		IRN(3)		IRN(4)		IRN(5)		IRN(6)	
	MSE	-SSIM	MSE	-SSIM	MSE	-SSIM	MSE	-SSIM	MSE	-SSIM	MSE	-SSIM
PSNR	27.4209	28.7636	27.9536	28.9173	28.4065	28.7902	28.9701	29.251	29.0032	29.6924	29.3527	31.1536
SSIM	0.9038	0.9781	0.9171	0.974	0.9202	0.9735	0.9461	0.9834	0.9543	0.9918	0.9651	0.9923
FADE	1.4116	1.296	1.4051	1.248	1.1883	1.1792	0.9632	0.836	0.8532	0.7966	0.8354	0.7318

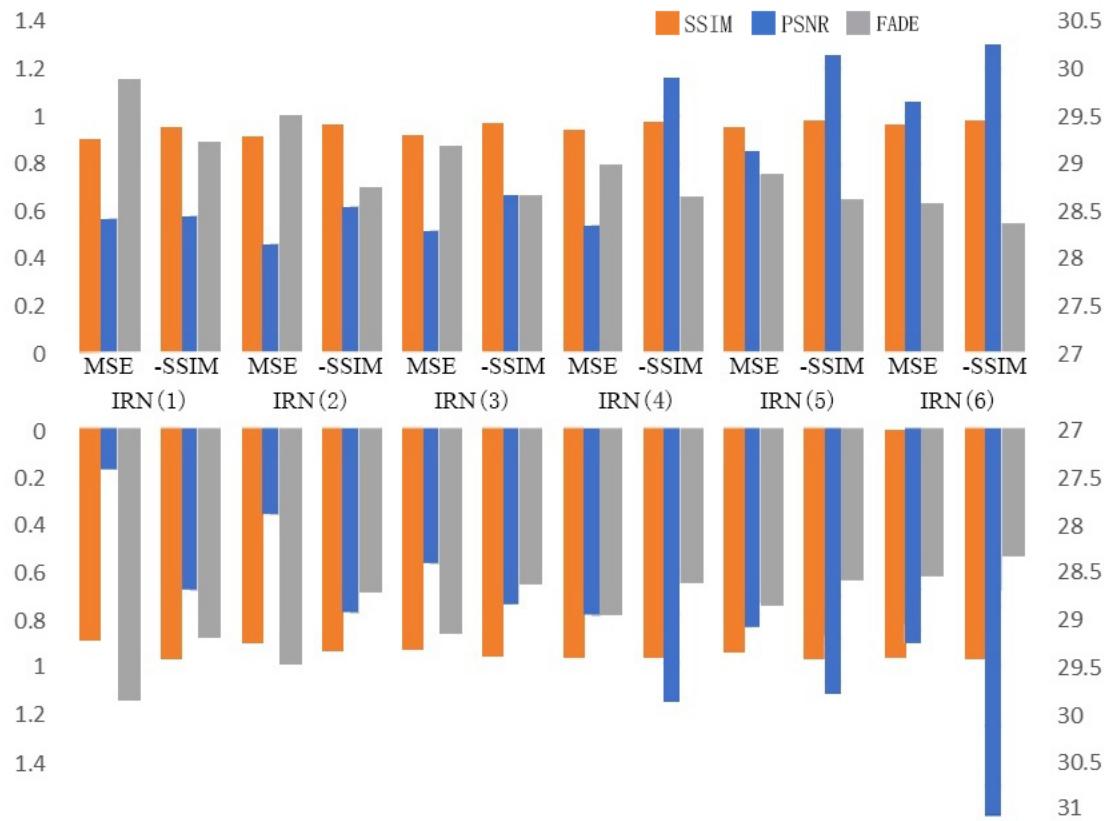


FIGURE 5: Intuitive comparison of objective indicators between MSE loss and negative SSIM loss under different iteration times. The upper part is the comparison in processing indoor composite images, and the lower part is the comparison in processing outdoor composite images. The higher the PSNR value the closer the SSIM value is to 1, the lower the FADE value, the higher the image quality.

6 is the best. By observing the results of model processing with different iterations, we find the processing results with the number of model iterations of 6-10 have little difference in visual perception, indicating that when the number of iterations is greater than 6, the model has tended to be saturated what means as the number of iterations increases, the model efficiency will decrease. So according to the objective

indicators and combined with subjective feelings, we believe that 6 times is the best number of iterations.

C. COMPARISONS WITH STATE-OF-THE-ART METHODS

Results on synthesis dataset. We tested the proposed network on RESIDE’s SOTS test set for objective and subjective evaluation, and compared it with advanced methods includ-

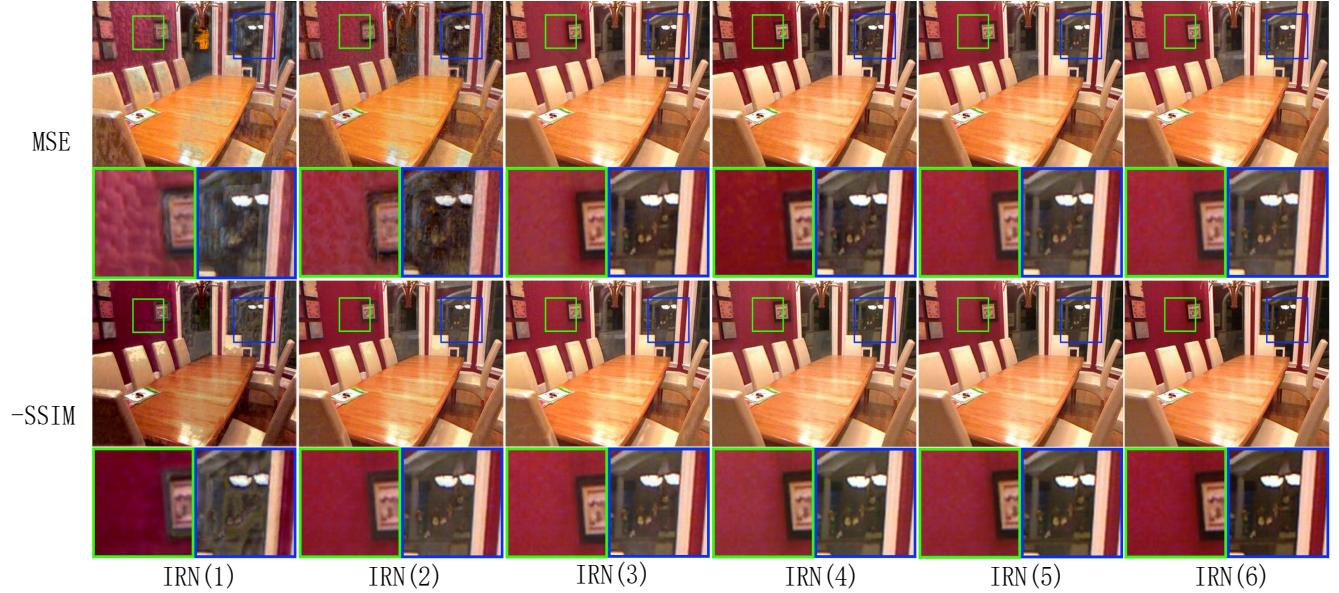


FIGURE 6: Under different iterations of the model, indoor synthetic images are used to train and test the difference between MSE loss and -SSIM loss in subjective visual perception. It is recommended to use a high-resolution monitor for viewing.

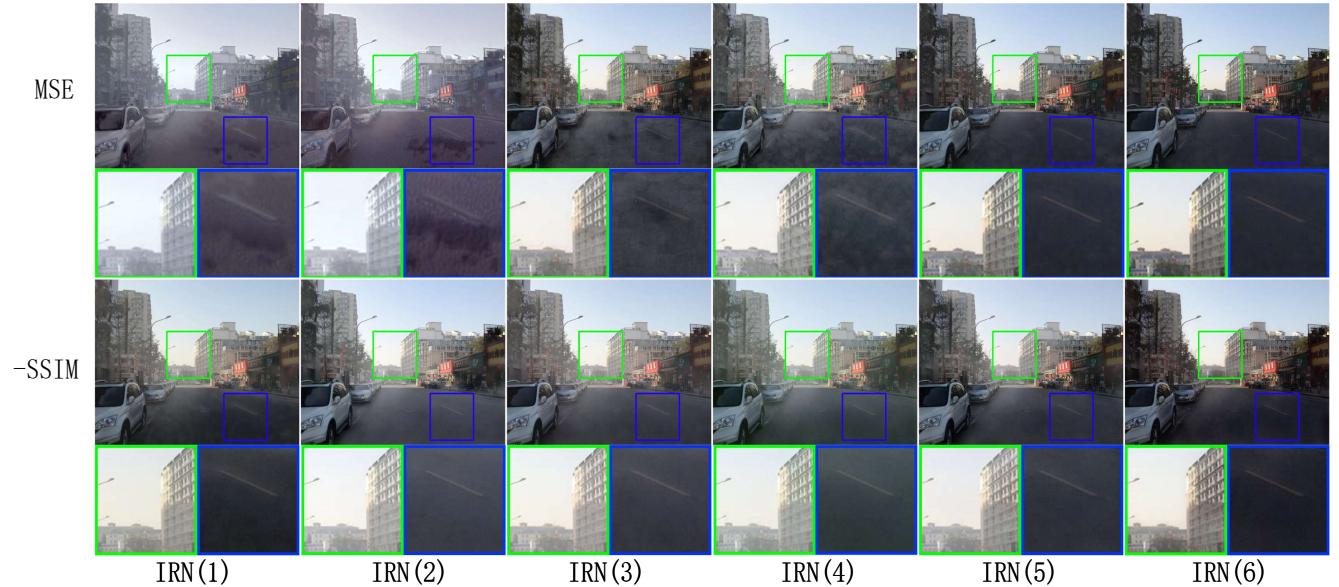


FIGURE 7: Under different iterations of the model, outdoor synthetic images are used to train and test the difference between MSE loss and -SSIM loss in subjective visual perception.

ing DCP, Hazy-lines, AOD-Net, GCANet, GridDehazeNet, FFA-Net and MSBDN. For fair comparison, we use the same data set to retrain the data-driven method, and use PSNR, SSIM, and FADE to quantitatively evaluate the output after dehazing. The value in TABLE 6 is the average value of the dehazing results of each method in the SOTS test set. The PSNR values of the proposed method in the SOTS indoor and outdoor test sets are ahead of other methods. SSIM and FADE are the same in all methods. At a high level, SSIM has a greater advantage over other methods in the outdoor test concentration.

Fig.12 shows the indoor composite image processing results of each method in SOTS. DCP has an obvious effect on removing smog, but there are problems with color distortion and aperture artifacts around objects. In the processing result of Hazy-lines, there is a serious color shift and some haze remains. In the processing results of AOD-Net, there are more haze residues. There is also a little haze remaining in the processing results of GCANet, for example, in the wall of the first picture, the wall color recovery is not good. In the processing result of GridBehazeNet, there is still residual haze, and black patches are prone to appear on the surface

TABLE 4: Use indoor synthetic images to test models with different iteration times, and compare them with PSNR, SSIM and FADE. The value in the table is the average of all images.

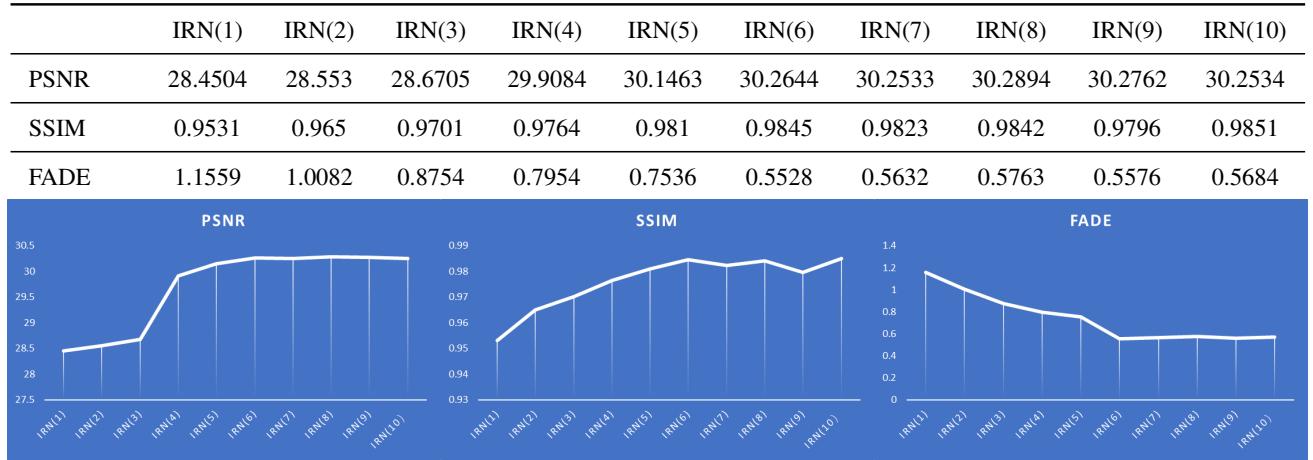


FIGURE 8: In the use of indoor synthetic images, the trend of objective evaluation indicators for models with different iteration times.

TABLE 5: Use outdoor synthetic images to test models with different iteration times, use PSNR, SSIM and FADE for comparison. The value in the table is the average of all images.

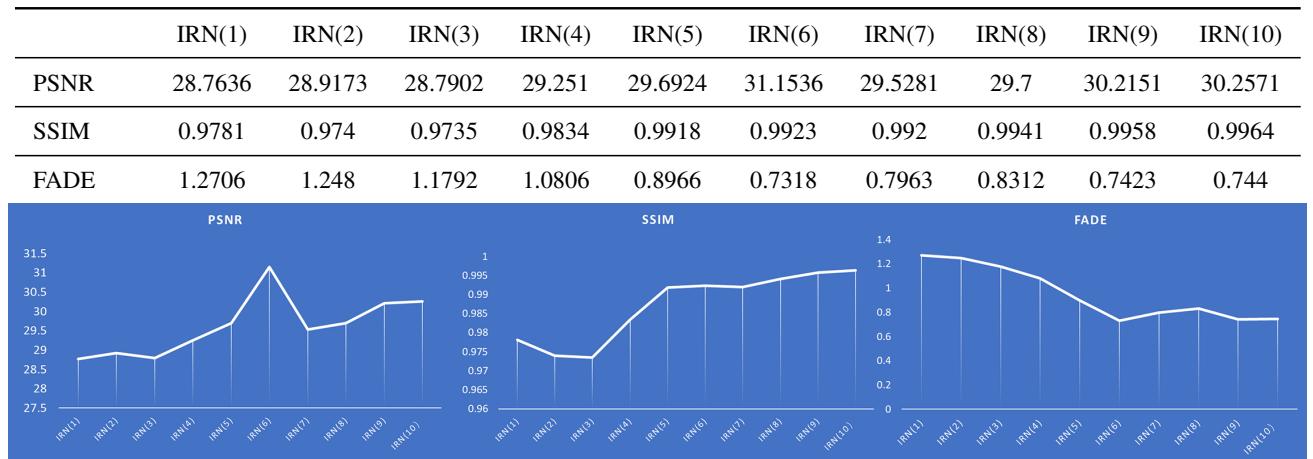


FIGURE 9: In the use of outdoor synthetic images, the trend of objective evaluation indicators for models with different iteration times.

TABLE 6: Use indoor synthetic images and outdoor synthetic images to test different models. The value in the table is the average of all images.

METHOD	Indoor			Outdoor		
	PSNR	SSIM	FADE	PSNR	SSIM	FADE
DCP	28.804	0.9803	0.5679	28.5063	0.9467	0.5613
Hazy-lines	28.1599	0.8157	0.4534	28.0783	0.8804	0.4338
AOD-Net	28.7184	0.9596	0.8052	28.2475	0.9701	0.7672
GCA-Net	28.5638	0.9667	0.7983	28.8829	0.9813	0.5941
GridDehazeNet	30.0035	0.9905	0.5499	28.5469	0.977	1.2656
FFA-Net	29.1523	0.9649	0.7943	30.1502	0.9676	0.5941
MSBDN	28.0113	0.9109	0.5364	27.9608	0.9465	0.8238
Ours	30.2644	0.9845	0.5528	31.1536	0.9923	0.7318

of the object. In the processing results of FFA-Net, there is more residual haze, and there are still color shifts. MSBDN

has a significant dehazing effect on indoor composite images, but in some pictures, there are still color shifts, such as the floor area in the third picture, and this method has darker and darker colors in areas with greater depth, the details are not obvious enough. Our method performs better in defogging indoor images, and the object material and details are retained to a greater extent. The color of the image after dehazing is closer to the original hazy-free image.

Fig.13 shows the outdoor composite image processing results of each method in SOTS. In the result of DCP processing, the color of the sky area is oversaturated and the visual effect is not good. In the processing result of Hazy-lines, the contrast of some pictures is too high, and there is color shift. In the result of AOD-Net processing, the color of the image is darker and darker, and a small amount of haze remains. In the

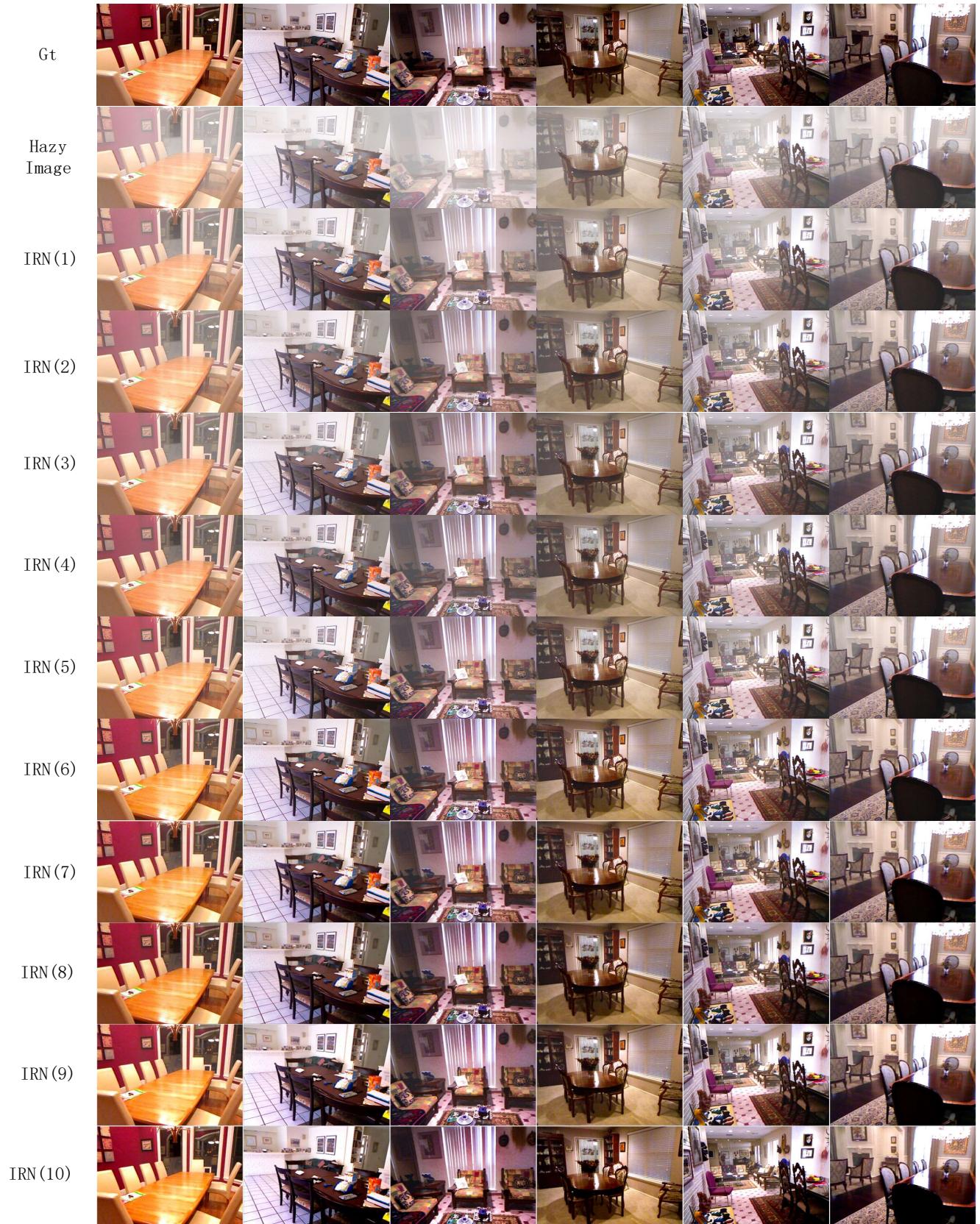


FIGURE 10: Subjective visual comparison of models with different iteration times in indoor synthetic images. It is recommended to use a high-resolution monitor for viewing.

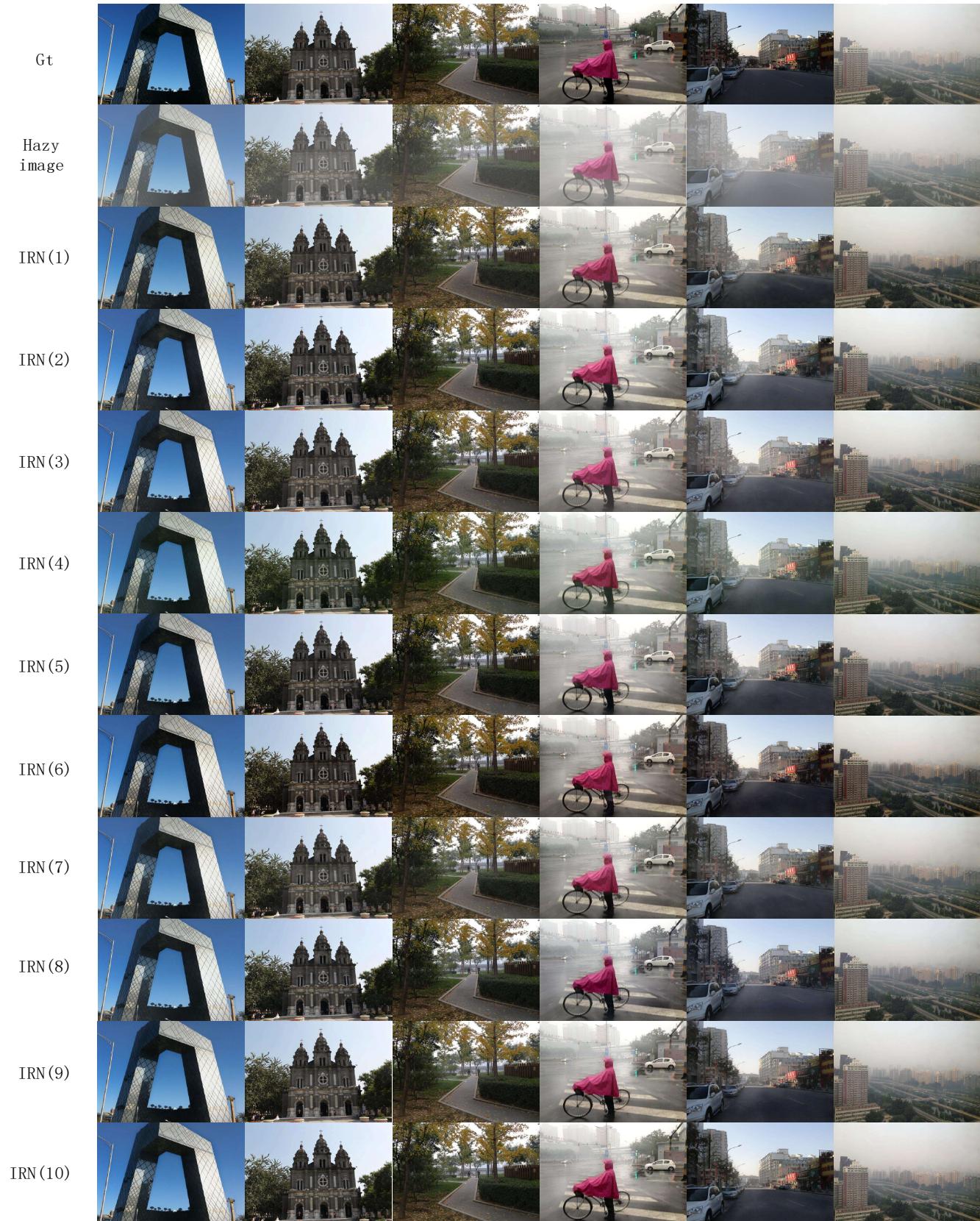


FIGURE 11: Subjective visual comparison of models with different iteration times in outdoor synthetic images.



FIGURE 12: Use indoor synthetic images to compare the subjective visual experience of different models.

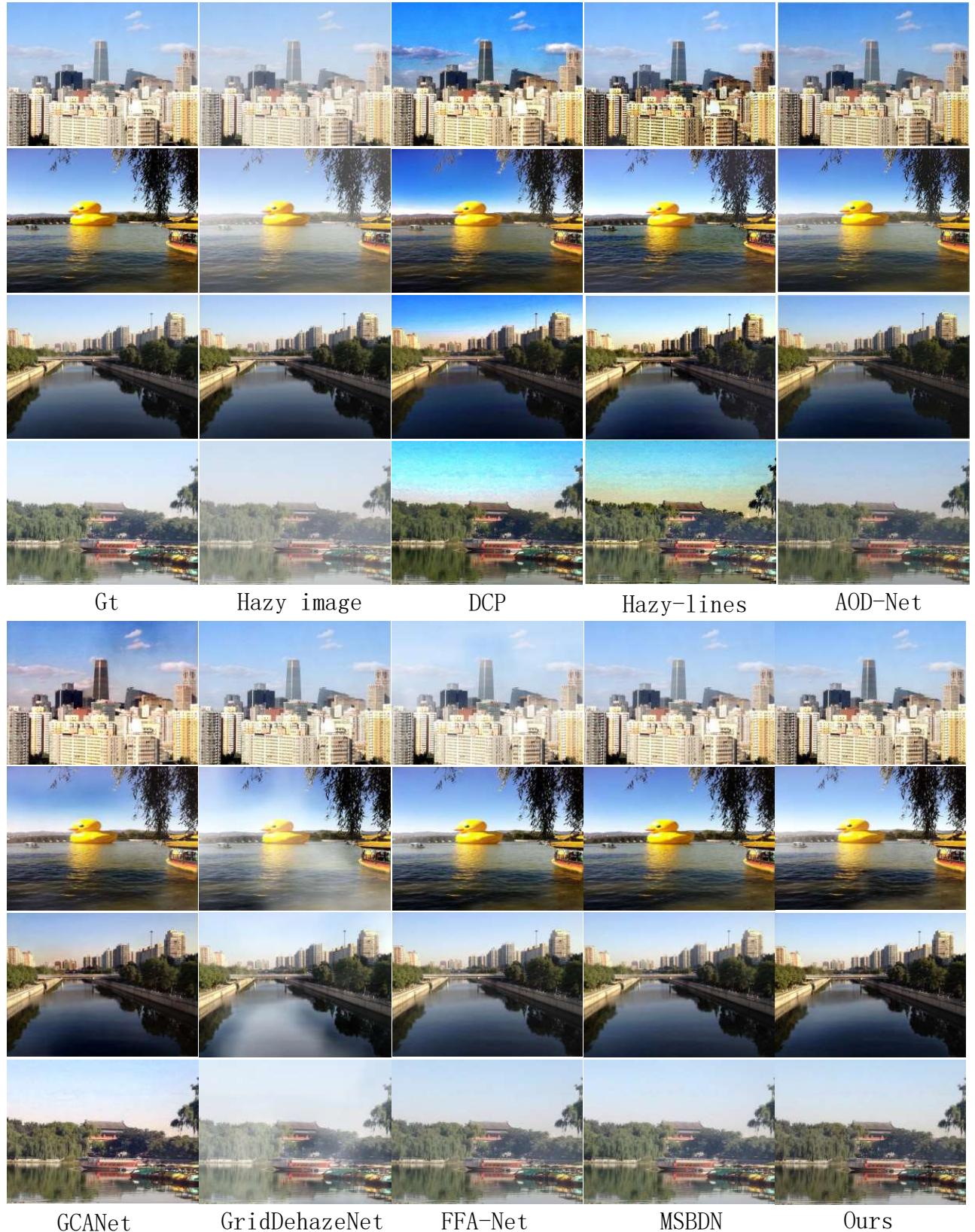


FIGURE 13: Use outdoor synthetic images to compare the subjective visual experience of different models.

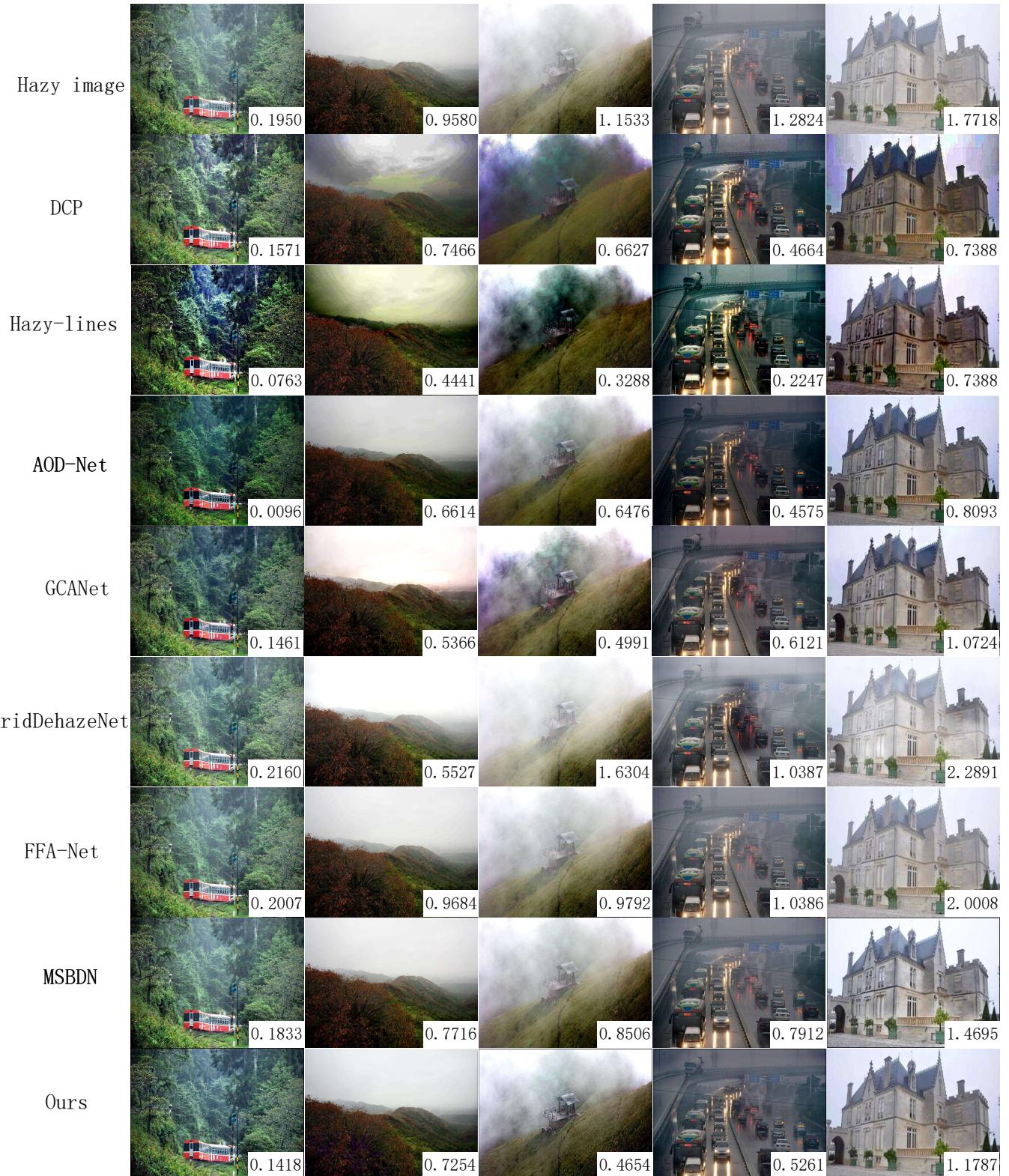


FIGURE 14: Different methods are used to process the comparison of haze images in real scenes. The value in the lower left corner of the image is the FADE value of the image. The smaller the FADE value, the lower the haze residue degree to a certain extent.

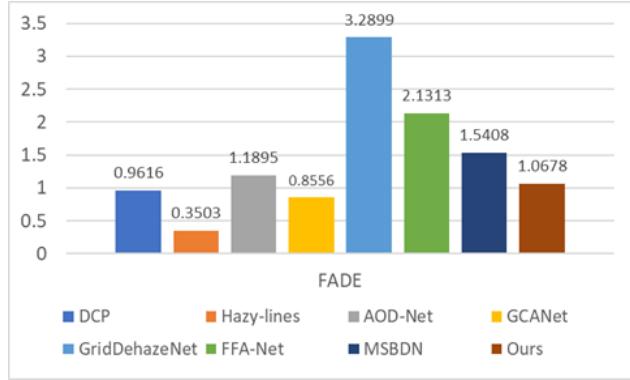


FIGURE 15: Use the test set of real scenes to test different methods. The different colors in the figure represent different methods, and the numbers in the figure represent the average FADE value of different methods. The smaller the value, the smaller the residual degree of haze.

processing result of GCANet, the sky area is not pure enough, a small amount of haze remains, and the contrast is high. In the result of GridDehazeNet processing, the image has a lot of haze and the overall contrast of the image is low. It can be clearly seen from the processing results of FFA-Net that in the sky area in the first picture, the haze is not well removed, and the color restoration of the tree area in the fourth picture is not obvious. In the MSBDN processing results, the haze removal effect is better, and individual images have low image contrast, such as the sky area in the first image, the buildings in the third image, and the buildings in the fourth image. In the processing results of our method, the haze can be effectively removed, the image as a whole is closer to the original hazy-free image, and the image has a better overall look and feel.

Based on the above, the comprehensive performance of the method proposed in this article is better than other methods in objective evaluation indicators, and it is also at a higher level in subjective visual perception. It is closer to the original clean image in terms of image color, structure, object surface texture and image details

Results on a real-world dataset. We use the real data set in RESIDE and the hazy image in the real scene collected by ourselves to test other methods and the methods mentioned in this article. The FADE in Fig.15 is the average value of the FADE of all test results. This article uses the FADE as An objective evaluation index for hazy image testing in real outdoor scenes.

Fig.14 shows the processing results of each method in the real scene. The five images represent five scenes: mist, medium density hazy, high density hazy, low light scene, outdoor building. The FADE of the current image is marked in the lower right corner of the image. Through observation, DCP works well in haze scenes, but in medium-concentration haze images, large areas of artifacts appear in the sky. Under high-concentration haze images, the color of the sky area

appears severely abnormal. In low-light scenes, hazy removal the effect is good, but there is partial image distortion. In the processing of outdoor buildings, the sky appears color shift and halo artifacts, and the building surface color is distorted. Hazy-lines achieved the lowest value in FADE's evaluation. Hazy-lines performed relatively well in haze scenes, but there would be a phenomenon of high contrast, resulting in unnatural image display. Hazy-lines also showed color shift and light the problem of halo artifacts, but compared to DCP has been reduced a lot, Hazy-lines has better dehazing effect, but the image restoration quality is lower.

ADO-Net has outstanding performance in haze images, and it has the best visual effects and FADE performance of all methods. However, in the medium-concentration haze image, the sky appears a little artifact, and the color is in the processing of high-concentration haze image In the low-light scene, although the FADE is low, the overall image is dark and the visual effect is not good. In the outdoor building scene, the color shift of the sky area and the surface of the building also appears, but it is relatively Great improvement in DCP. GCANet is the more advanced method among the existing methods, and the performance of FADE is also very good. The obvious disadvantage is that the color shift occurs in the dense fog scene and some noise appears in the picture. GridDehazeNet's performance in these five images is not outstanding enough, indicating that this method is not capable of dealing with real scenes, but the processed image does not show sky artifacts and color shifts. The performance of FFA-Net in real scenes is inferior to the performance when processing composite images. In the processing results of FFA-Net, it can be seen that whether it is mist or dense fog, there is still more haze in the image, and some images the FADE value is even higher than the original image. MSBDN has a certain effect when processing haze images of different concentrations. In low-light scenes, the overall image performance is slightly darker. In the performance of outdoor buildings, the color of the building surface is uneven. The performance of our method in low-concentration hazy is at a relatively high level among all methods. There is no sky artifact halo in the medium-concentration haze image, and the best performance in the high-concentration haze image, although it does not remove all the haze in the image, but the clarity of the objects in the image is the highest of all methods, and FADE is the lowest value. In low-light scenes, the FADE is at a higher level among all methods, compared to MSBDN. Processing results, the processing results of the method proposed in this paper are clearer. In the processing of outdoor architectural images, the color of the building surface is more pure, and it performs best in all images.

Although the method proposed in this paper uses synthetic data sets for training, it still achieves a relatively ideal effect in real scenarios, and its dehazing effect is better than other methods in terms of synthesis.

V. CONCLUSION

In this paper, we propose an Iterative Residual Network for image dehazing. This method does not rely on atmospheric physical models and directly generates clean images in an end-to-end manner. By designing an iterable computing unit and corresponding model structure, the number of network parameters is greatly reduced, making the network training faster. Through experiments, it is found that by increasing the number of iterations, the dehazing ability of the model will increase. At the same time, in the iterative calculation process of the model, in order to prevent the loss of image details under high iteration times, we add to the Dehaze Unit An LSTM unit is used to process the features at different iterations to preserve image features. When the model extracts deep-level features, multiple residual units are combined to ensure the stability of network training. After a large number of experimental comparisons, the superiority of the method proposed in this paper is proved in synthetic images and real scenes.

ACKNOWLEDGMENT

The authors acknowledge the National Natural Science Foundation of China (Grant nos. 61772319, 61976125, 61976124 and 61773244), and Shandong Natural Science Foundation of China (Grant no. ZR2017MF049).

REFERENCES

- [1] E. J. McCartney, "Optics of the atmosphere: Scattering by molecules and particles," *International journal of computer vision*, vol. 28, no. 11, pp. 521–521, 1977.
- [2] S. G. Narasimhan and S. K. Nayar, "Chromatic framework for vision in bad weather," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662)*, vol. 1. IEEE, 2000, pp. 598–605.
- [3] S. G. Narasimhan and S. K. Nayar, "Vision and the atmosphere," *International journal of computer vision*, vol. 48, no. 3, pp. 233–254, 2002.
- [4] K. He, J. Sun, and X. 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.
- [5] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [7] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "Dehazenet: An end-to-end system for single image haze removal," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5187–5198, 2016.
- [8] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "Aod-net: All-in-one dehazing network," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 4770–4778.
- [9] Y. Qu, Y. Chen, J. Huang, and Y. Xie, "Enhanced pix2pix dehazing network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 8160–8168.
- [10] X. Liu, Y. Ma, Z. Shi, and J. Chen, "Griddehazenet: Attention-based multi-scale network for image dehazing," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 7314–7323.
- [11] H. Dong, J. Pan, L. Xiang, Z. Hu, X. Zhang, F. Wang, and 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.
- [12] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [13] J. Li, G. Li, and H. Fan, "Image dehazing using residual-based deep cnn," *IEEE Access*, vol. 6, pp. 26831–26842, 2018.
- [14] R. Yasrala, V. A. Sindagi, and V. M. Patel, "Syn2real transfer learning for image deraining using gaussian processes," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 2726–2736.
- [15] K. Zhang, W. Luo, Y. Zhong, L. Ma, B. Stenger, W. Liu, and H. Li, "Deblurring by realistic blurring," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 2737–2746.
- [16] W. Zhou, A. C. Bovik, H. R. Sheikh, and P. Eero, "Simonselli. image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.
- [17] L. K. Choi, J. You, and A. C. Bovik, "Referenceless prediction of perceptual fog density and perceptual image defogging," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3888–3901, 2015.
- [18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [19] L. R. Bissonnette, "Imaging through fog and rain," *Optical Engineering*, vol. 31, no. 5, pp. 1045–1053, 1992.
- [20] J. P. Oakley and B. L. Satherley, "Improving image quality in poor visibility conditions using a physical model for contrast degradation," *IEEE transactions on image processing*, vol. 7, no. 2, pp. 167–179, 1998.
- [21] S. G. Narasimhan and S. K. Nayar, "Contrast restoration of weather degraded images," *IEEE transactions on pattern analysis and machine intelligence*, vol. 25, no. 6, pp. 713–724, 2003.
- [22] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, "Polarization-based vision through haze," *Applied optics*, vol. 42, no. 3, pp. 511–525, 2003.
- [23] R. Fattal, "Single image dehazing," *ACM transactions on graphics (TOG)*, vol. 27, no. 3, pp. 1–9, 2008.
- [24] R. T. Tan, "Visibility in bad weather from a single image," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2008, pp. 1–8.
- [25] 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.
- [26] L. Shen, Y. Zhao, Q. Peng, J. C.-W. Chan, and S. G. Kong, "An iterative image dehazing method with polarization," *IEEE Transactions on Multimedia*, vol. 21, no. 5, pp. 1093–1107, 2018.
- [27] T. Pohlen, A. Hermans, M. Mathias, and B. Leibe, "Full-resolution residual networks for semantic segmentation in street scenes," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4151–4160.
- [28] Y. Song, C. Ma, L. Gong, J. Zhang, R. W. Lau, and M.-H. Yang, "Crest: Convolutional residual learning for visual tracking," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2555–2564.
- [29] C. Guo, C. Li, J. Guo, R. Cong, H. Fu, and P. Han, "Hierarchical features driven residual learning for depth map super-resolution," *IEEE Transactions on Image Processing*, vol. 28, no. 5, pp. 2545–2557, 2018.
- [30] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, "Single image dehazing via multi-scale convolutional neural networks," in *European conference on computer vision*. Springer, 2016, pp. 154–169.
- [31] H. Zhang and V. M. Patel, "Densely connected pyramid dehazing network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 3194–3203.
- [32] K. Mei, A. Jiang, J. Li, and M. Wang, "Progressive feature fusion network for realistic image dehazing," in *Asian conference on computer vision*. Springer, 2018, pp. 203–215.
- [33] N. BharathRaj and N. Venkateswaran, "Single image haze removal using a generative adversarial network," *ArXiv*, vol. abs/1810.09479, 2018.
- [34] D. Chen, M. He, Q. Fan, J. Liao, L. Zhang, D. Hou, L. Yuan, and 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.
- [35] A. Dudhane, H. Aulakh, and S. Murala, "Ri-gan: An end-to-end network for single image haze removal," *06 2019*, pp. 2014–2023.
- [36] D. Engin, A. Genç, and H. Kemal Ekenel, "Cycle-dehaze: Enhanced cycle-gan for single image dehazing," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 825–833.
- [37] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Jia, "Ffa-net: Feature fusion attention network for single image dehazing," in *AAAI*, 2020, pp. 11908–11915.

- [38] B. R. N. and V. N., "Single image haze removal using a generative adversarial network," 2018.
- [39] W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W. Liu, and 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.
- [40] J. Kim, J. Kwon Lee, and K. Mu Lee, "Deeply-recursive convolutional network for image super-resolution," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 1637–1645.
- [41] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735–1780, 1997.
- [42] Z. Fan, H. Wu, X. Fu, Y. Hunag, and X. Ding, "Residual-guide feature fusion network for single image deraining," arXiv preprint arXiv:1804.07493, 2018.
- [43] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in Proceedings of the IEEE international conference on computer vision, 2017, pp. 2223–2232.
- [44] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training gans," in Advances in neural information processing systems, 2016, pp. 2234–2242.
- [45] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in European conference on computer vision. Springer, 2016, pp. 694–711.
- [46] M.-F. Tsai, T.-Y. Liu, T. Qin, H.-H. Chen, and W.-Y. Ma, "Frank: a ranking method with fidelity loss," in Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, 2007, pp. 383–390.
- [47] D. Ren, W. Zuo, Q. Hu, P. Zhu, and D. Meng, "Progressive image deraining networks: A better and simpler baseline," pp. 3937–3946, 2019.
- [48] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang, "Reside: A benchmark for single image dehazing," arXiv preprint arXiv:1712.04143, vol. 1, 2017.
- [49] Y. Zhang, S. Song, E. Yumer, M. Savva, J.-Y. Lee, H. Jin, and T. Funkhouser, "Physically-based rendering for indoor scene understanding using convolutional neural networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 5287–5295.



JINJIANG LI received the B. S. and M. S. degrees in computer science from Taiyuan University of Technology, Taiyuan, China, in 2001 and 2004, respectively, the Ph. D. degree in computer science from Shandong University, Jinan, China, in 2010. From 2004 to 2006, he was an assistant research fellow at the institute of computer science and technology of Peking University, Beijing, China. From 2012 to 2014, he was a Post-Doctoral Fellow at Tsinghua University, Beijing, China. He is currently a Professor at the school of computer science and technology, Shandong Technology and Business University. His research interests include image processing, computer graphics, computer vision, and machine learning.

• • •



ZHEN HUA received the B.S. and M.S. degrees in electrical automation from Taiyuan University of Technology, Taiyuan, China, in 1989 and 1992, respectively, the Ph.D. degree in electronic information engineering from China University of Mining and Technology, Beijing, China, in 2008. She is currently a professor at Shandong Technology and Business University. Her research interests include computer aided geometric design, information visualization, virtual reality, and image processing.



GUODONG FAN received his bachelor's degree from the School of Computer Science and Technology, Shandong Technology and Business University, Yantai, China in 2018. Currently studying for a master's degree in the School of Information and Electronic Engineering, Shandong Technology and Business University, Yantai, Shandong. His research interests include computer graphics, computer vision, and image processing.