IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 8 June 2022, accepted 19 June 2022, date of publication 23 June 2022, date of current version 1 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3185627

Remote Sensing Image Dehazing Based on an Attention Convolutional Neural Network

ZHIJIE HE[1,2,3], CAILAN GONG[1,2], YONG HU[1,2], AND LAN LI[2,3]

[1]Key Laboratory of Infrared System Detection and Imaging Technology, Chinese Academy of Sciences, Shanghai 200083, China
[2]Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China
[3]Shanghai Institute of Technical Physics, University of Chinese Academy of Sciences, Beijing 100049, China

Corresponding author: Cailan Gong (gcl@mail.sitp.ac.cn)

ABSTRACT Haze may affect the quality of optical remote sensing images, thus limiting the scope of their application. Remote sensing image dehazing has become important in remote sensing image preprocessing, promoting the use of remote sensing data and the precision of target recognition. Existing remote sensing dehazing methods based on simplified atmospheric degradation models are not suitable for the removal of heterogeneous haze that exist in remote sensing images. For this purpose, this study proposes an end-to-end convolutional neural network based on attention mechanism, in which the residual block structure combines both channel and spatial attention mechanisms, and establishes a synthetic high-resolution haze image dataset for full training. Thus, it obtains the desired dehazing model. Finally, this study investigates the dehazing model using a GF-1 image and compares it with existing dehazing methods. The results show that the proposed method improved the image similarity, color authenticity, and haze residue level.

INDEX TERMS Attention mechanism, convolutional neural network, dehazing, remote sensing image.

I. INTRODUCTION

With the rapid development of new multi-platform sensor technology, continuously improved remote sensing image data acquisition capability has become crucial for fields such as land resources, environmental protection, and urban planning [1]. However, in optical remote sensing satellites imaging, significant blur and low contrast may occur in remote sensing images owing to ice, moisture, dust, and other particles in the atmosphere; this significantly affects subsequent applications. Therefore, effective remote sensing image dehazing is crucial in improving the use of existing remote sensing images as a preprocessing technique.

A dehazing method based on the message complementary sense among multiple images for multi-temporal data in remote sensing images is available. However, compared with single image dehazing methods, this method needs auxiliary data, especially for high-resolution remote sensing data with low temporal resolution; it is poor in applicability. Therefore, single image dehazing methods have attracted attention in the academic community. Generally, there are two types of conventional single image dehazing methods: (1) single-image dehazing based on image enhancement, usually processes histograms [2] and enhances the contrast [3] and saturation [4] of images. Moreover, dehazing methods are based on homomorphic filtering [5] and retinex color invariance theory [6]. In recent years, studies that fuse different exposure images to effectively maintain the color of the image while dehazing and restoring local details [7], [8] have been conducted. (2) Single-image dehazing, which is based on a scattering model, involves the establishment of the image degradation model for hazy days, based on the formation mechanism of images disturbed by hazes [9]–[12]. Here, the parameters of the model are solved based on prior knowledge and the images are restored. Based on the assumption of constant local albedo, an image degradation model was solved for hazy days using independent component analysis [9]. He *et al.* [10] proposed a dark channel prior (DCP) method based on observations and statistics for numerous images on hazy days for solving the atmospheric scattering model on hazy days, which achieved good results. Since then, several dehazing methods have been modified based on the dark-channel prior method. The methods above are based on prior knowledge of image dehazing, which requires manual extraction of image features. This presents some limitations associated with processing images of different scenes.

In recent years, with the rapid development of deep learning and neural network technology in the fields of computer

The associate editor coordinating the review of this manuscript and approving it for publication was Oguzhan Urhan.

vision and image processing, several learning-based methods have emerged in the field of image dehazing. Initially, some researchers estimated the transmittance in the image degradation model for hazy days by constructing neural networks and then restoring the hazy images based on the model. Cai *et al.* [13] adopted a multi-scale convolutional neural network (CNN) with hazy images as the input to output the corresponding transmittance diagrams. The images were restored using a degradation model. Ren *et al.* [14] first input the hazy images into a coarse-scale network to obtain a rough transmittance diagram. Subsequently, they input the hazy images into a fine-scale network to obtain the optimized transmittance diagram; finally, the images were further refined. The dehazing method based on the transmittance estimation network achieved a good effect; however, this was dependent on the atmospheric scattering model. The atmospheric scattering model is only a simplified simulation of hazy imaging; therefore, these methods may not be sufficient for image feature extraction. In some cases, the method sets the atmospheric light value as a global constant, leading to color distortion after dehazing and insufficient dehazing of the images. Thus, compared with conventional methods such as the DCP method, this method has no clear advantages. The direct end-to-end image restoration network has achieved a dehazing effect finer than that of previous methods for its powerful neural network capacity for feature extraction. Li, *et al.* [15] transformed the degradation model for hazy days and proposed a direct end-to-end dehazing neural network based on the transformed model, which could obtain haze-free images directly, omitting the procedure for estimating intermediate parameters. Chen *et al.* [16]proposed an end-to-end gated context aggregation network that combined smooth cavity convolution and multilevel feature fusion to obtain a finer dehazing effect. Yu *et al.* [17] proposed a fully end-to-end generative adversarial network with a fusion discriminator (FD-GAN) for image dehazing. For the FD-GAN, both low- and high-frequency components were used as the discriminator to further distinguish the hazy image from the haze-free image. This guides the generator to output more natural and realistic haze-free images. With the increasing number of natural image datasets for hazy days, these methods achieved good results after extensive training. Inspired by the application of deep learning for natural image dehazing, multiple studies using neural networks for remote sensing image dehazing have emerged. Enomoto *et al.* [18] used near-infrared images as auxiliary data along with RGB images as input to a conditional generation network to generate haze-free visible light images. Jiang *et al.* [19] used a multi-scale convolutional neural network structure similar to [13], while adding residual blocks (RBs) to the network, which was used to restore hazy images of Landsat 8 Operational Land Imager (OLI). Qin *et al.* [20] used a network structure that connects multiple residual modules in parallel to remove haze from multiple-channel remote sensing images. Mehta *et al.* [21] proposed a network module named Sky-GAN for haze removal in aerial images, which included a hazy-to-hyperspectral (H2H) module and a conditional generative adversarial network (cGAN) module. The H2H module reconstructed multispectral images from RGB bands to expand hazy aerial image datasets whereas the cGAN module used the reconstructed multispectral data to remove haze. Remote sensing images often have rich background information and uneven haze distribution. In a multiscale structure for general dehazing networks, one scale influences another in terms of precision. Considering that multi-scale estimation is usually conducted in a continuous manner, the adaptability of the network model may be affected when applied to remote sensing images with more complex feature information. The network may not fully utilize the information of remote sensing images and the dehazing effect on heterogeneous haze, widely present in remote sensing images, may be poor. Meanwhile, unlike natural image dehazing, there is no large-scale publicly available dataset for remote sensing image dehazing. The applicability of the dehazing model may be insufficient when applied to remote sensing images of different scales.

To this end, this study proposes an end-to-end single image dehazing network based on an attention mechanism. This network sets up the attention module in both spatial and channel dimensions with a basic encoder-decoder structure using residual dense blocks to sufficiently extract image features and effectively remove heterogeneous haze. A high resolution hazy image dataset based on GF-1 data was created, on which the dehazing model was applied.

II. METHODS

Significant success has been achieved in visual tasks such as image super-resolution [22], using the encoder–decoder structure. Based on this structure, this study proposed a coding–decoding model combining the RB and attention mechanism. First, the model transformed the input hazy images into a high-dimensional feature images through the coding part and performed feature extraction through the RB, channel attention (CA), and spatial attention (SA) modules. Second, the feature map was decoded back to the original image space to obtain the haze feature residue. Third, dehazed images were obtained after the addition of the input images. Fig. 1 shows the overall network structure. Details of the modules in the network structure are given in Table 1.

A. ENCODING–DECODING MODULE

The encoder consisted of two convolutional blocks, each composed of a convolutional layer, batch normalization, and ReLU. The encoder down sampled the input images to reduce their dimensions and increase the number of feature channels; the decoder was partially similar to the encoder in structure (see Fig. 1). The output convolutional layer was changed into a deconvolutional layer for up sampling and the feature information was restored to the size of the original image.

B. RESIDUAL BLOCK

The training and convergence of the network depended on the reverse transmission of training errors. An excessive network

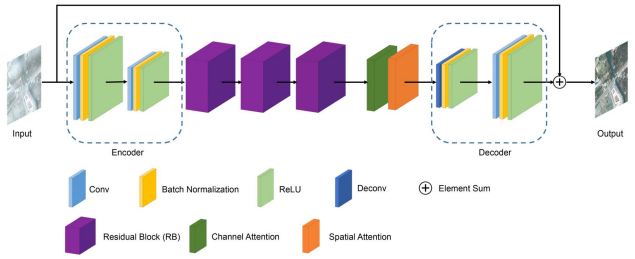

FIGURE 1. Overall network structure.

TABLE 1. Details of the network.

Module	Layer	Operation	Output size
input	-	-	3×512×512
encoder	Conv	Kernel 7×7; BN; ReLu	32×256×256
	Conv	Kernel 5×5; BN; ReLu	128×128×128
RB-1	Conv	Kernel 5×5; IN; ReLu	256×64×64
	Conv	Kernel 3×3; IN; ReLu	512×32×32
RB-2	Conv	Kernel 3×3; IN; ReLu	512×32×32
	Conv	Kernel 3×3; IN; ReLu	512×32×32
RB-3	Deconv	Kernel 3×3; IN; ReLu	256×64×64
	Deconv	Kernel 3×3; IN; ReLu	128×128×128
CA	Global Avg pool Global Max pool		128×128×128
	Conv	(Kernel 5×5)×2; ReLu×2	
SA	Conv	(Kernel 5×5)×2; ReLu	128×128×128
decoder	Deconv	Kernel 5×5; BN; ReLu	32×256×256
	Deconv	Kernel 7×7; BN; ReLu	3×512×512

depth might eliminate the gradient, leading to the nonconvergence of the network. This study performed feature learning for the feature map of the encoder with an RB structure consisting of two convolutional groups (see Fig. 2). Each convolutional group consisted of convolutional, instance normalization, and ReLU layers. Cross-layer elements were added to further focus the network on effective information and reduce the influence of information from low-haze or low-frequency regions.

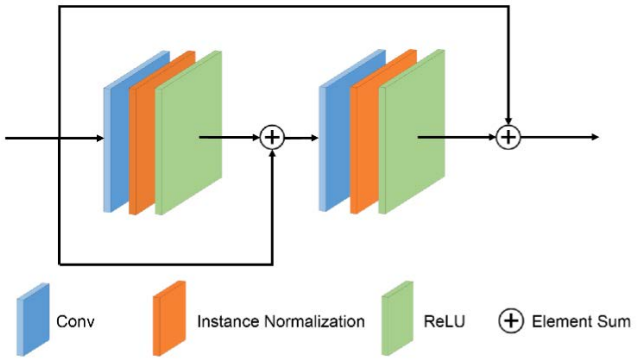

FIGURE 2. Structure of RB.

C. ATTENTION MECHANISM MODULE

For heterogeneous haze on satellite images, information on different haze thicknesses had to be effectively distinguished and processed. However, the conventional CNN dehazing models treated the features of channels and spatial dimensions equally and it was difficult for them to process heterogeneous haze in images. Therefore, this study introduced an attention mechanism to the network. Similar to human visual attention, the attention mechanism screened out the information that was most important for the current task from numerous pieces of information. It became critical to improve the network performance because the attention mechanism was widely used in the neural network. Therefore, this study introduced the channel and space attention modules to the model. Feature attention modules could selectively process or treat different channels and spatial pixels and generate different weights for channels and spatial pixel features, thus advancing the feature expression ability of the network.

The CA module extracted feature differences between channels to effectively remove haze from images. First, the module aggregated spatial information in the channel through global average pooling and global maximum pooling. For global pooling operations, see (1) and (2):

$$g_c = H_{\text{avg}}(F_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} X_c(i,j) \quad (1)$$

$$m_c = H_{\max}(F_c) = \max_{(i,j) \in X_c} X_c(i,j) \quad (2)$$

where g_c and m_c represent global average pooling and global maximum pooling, respectively, and $X_c(i,j)$ represents the value of $X_c$in Channel c at (i,j). Global pooling changed the size of the feature map from $C \times H \times W$ to $C \times 1 \times 1$. Different weights for different channels were obtained through global pooling. Subsequently, the feature maps activated functions through two convolutional layers with sigmoid and ReLU, respectively, and the functions were added as follows.

$$A_c = \sigma(\text{Conv}(\delta(\text{Conv}(g_c))) + \text{Conv}(\delta(\text{Conv}(m_c)))) \quad (3)$$

where σ is the sigmoid function, δ is the ReLU function, A_c is the weight at the channel level, and Conv is the convolution operation. The output of CA was obtained by multiplying the corresponding elements in sequence with input F_c with weight A_c.

$$F_c^* = A_c \otimes F_c \tag{4}$$

Thus, the CA module converted the global spatial information of the channel into a channel descriptor, i.e., F_c^*. Fig. 3 shows a schematic of the CA module. All the convolution operations performed were of 1×1 convolution after pooling. The features obtained by this point-by-point convolution calculation integrated the information from all channels but did not mix the information across space; this was conducive to distinguishing channel feature learning from spatial feature learning.

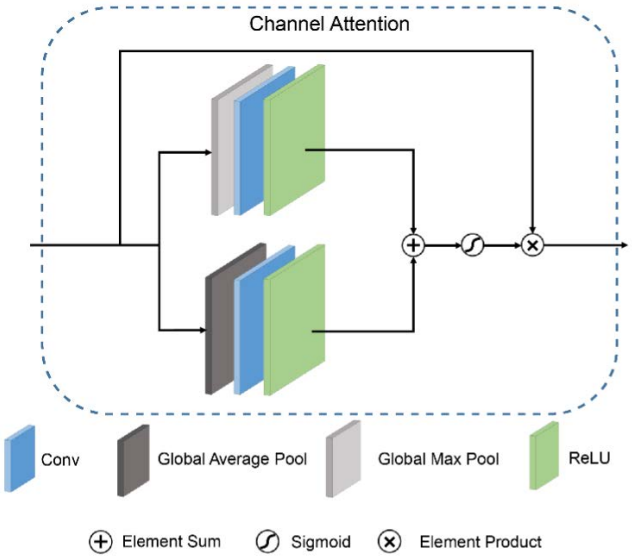

FIGURE 3. CA module.

Similar to CA, the SA module effectively dealt with the uneven distribution of haze at different image pixels such that the network paid more attention to the features of high-haze and high-frequency regions. The channel descriptor generated by the CA module, i.e., F_c^*, changed the size of the feature map from $C \times H \times W$ to $1 \times H \times W$ through the convolutional layer and activated the function and convolutional layer; see (5).

$$A_s = \sigma\left(\text{Conv}\left(\delta\left(\text{Conv}\left(F_c^*\right)\right)\right)\right) \tag{5}$$

where σ represents the sigmoid function, δ is the ReLU function, and A_s represents the weight of the space range. Finally, F_c^* and A_s were multiplied by the corresponding elements.

$$\tilde{F} = A_s F_c^* \tag{6}$$

where $\tilde{F}$ is the output for the entire feature attention module. Fig. 4 is a schematic of the SA module.

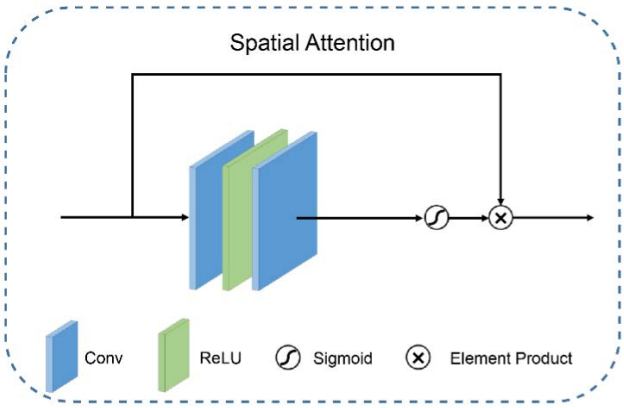

FIGURE 4. SA module.

D. LOSS FUNCTION

The selection and setting of the loss function were critical for training the entire neural network. This study used loss L_1 combined with perceptual loss [23] as a loss function. Loss L_1 provided a quantitative measure of the difference between hazy and haze-free images. Because norm L_1 could prevent potential gradient explosion, its sensitivity to outliers was lower than that of the mean square error loss.

In Loss L_1, $\hat{J}_i(x)$ represents the intensity of color channel i of pixel x in the dehazing image, $J_i(x)$ is the truth value image corresponding to $\hat{J}_i(x)$, and L_1 is the total number of pixels. The Loss L_1 is expressed as follows:

$$L_1 = \frac{1}{N}\sum_{x=1}^{N}\sum_{i=1}^{3} F_S\left(\hat{J}_i(x) - J_i(x)\right) \tag{7}$$

where

$$F_S(e) = \begin{cases} 0.5e^2, & \text{if } |e| < 1 \\ |e| - 0.5, & \text{otherwise} \end{cases} \tag{8}$$

In contrast to per-pixel loss, perceptual loss used multi-scale features extracted from a pre-trained deep neural network to quantitatively estimate visual differences between the image and the real situation on the ground. This study adopted VGG16 [24] pre-trained on ImageNet [25] as a loss network and extracted features from the last layer of the first three phases (i.e., Conv1-2, Conv2-2, and Conv3-3). Perceptual loss is defined as:

$$L_P = \sum_{j=1}^{3} \frac{1}{C_j H_j W_j} \|\phi_j(\hat{J}) - \phi_j(J)\|_2^2 \tag{9}$$

where $\phi_j(\hat{J})$, $\left(\phi_j(J)\right), j = 1, 2, 3$ represent the above three VGG16 feature maps associated with dehazing image $\hat{J}$ (truth image: J) and $C_j H_j W_j$ are the dimensions of the feature map $\phi_j(\hat{J})$, $\left(\phi_j(J)\right)$.

Finally, the study takes Loss L_1 combined with perceptual loss as the final loss function for this experiment; see (10).

$$L = L_S + \omega L_P \tag{10}$$

FIGURE 5. Synthetic haze remote sensing image dataset.

where ω is a parameter used to adjust the relative weights of the two loss components. Here, ω was set to 0.03.

III. EXPERIMENT

This section describes the construction of a dataset, implementation of training details, and quantitative evaluation of the dehazing effect based on Gf-1 remote sensing images with haze.

A. DATA SET CONSTRUCTION

Normally, it would be difficult to obtain many remote sensing images with haze and the corresponding haze-free images. Differences in ground information are often caused by factors such as weather and light. Therefore, this study synthesized hazy images to construct a training set for a learning method that required a large amount of training data. Because haze in remote sensing images is often of heterogeneous distribution, the commonly used construction method based on the degradation model [13] was not suitable for training sets of hazy remote sensing images. Berlin noise [26] was often used to simulate the texture in nature, and it was similar to the haze in remote sensing images in distribution. Therefore, this study used Berlin noise and haze-free remote sensing images to synthesize hazy images as training data. Here, 10 GF-1\MSS multispectral images with 4,548 $\times$ 4,544 dimensions, acquired in sunny weather in Shanghai and its surrounding areas from 2016 to 2019, were selected. Considering that appropriate dimensions were required for the training data, the images were cut to 512 $\times$ 512 pixels, Berlin noise was added to synthesize the training data. For each GF-1 image, the RGB band was selected and cut into 100 512 $\times$ 512 images (including certain overlapping edges). Through image augmentation and the addition of Berlin noise with different features, 11,035 sets of training data were synthesized. Fig. 5 shows some of these results. Out of the training data, 90% and 10% were used as the training and test sets, respectively.

B. DETAILS FOR TRAINING

Here, the method was realized using PyTorch, and the model training was completed on an NVIDIA 1080Ti GPU. During the training, the images were processed in random crop mode, and the dimensions of the image blocks that were inputs for the model were 512 $\times$ 512. The Adam optimizer with a batch size of 12 was used to train the model, and 0.9 and 0.999 were adopted as the default values for the momentum parameters $\beta 1$ and $\beta 2$, respectively. One hundred iterations were performed with 0.001 as the initial learning rate of the model; the learning rate decreased to half of the previous rate after every 20 iterations.

C. EXPERIMENTAL RESULTS AND DISCUSSION

In this experiment, synthetic GF-1\MSS hazy images were assessed and compared with those from the classical DCP method and those from the neural-network methods of DehazeNet and aerosol optical thickness (AOD)-NET. Fig. 6 shows the experimental results. Visually, DCP, DehazeNet, and AODNet methods restored hazy images to a certain extent; however, compared with the method proposed, the three methods had visible haze residues and color bias. Moreover, for large gray areas covered by haze in the images, the three methods had more haze residues (See the dehazing effect of line 2). These results were analyzed in terms of commonly used image-restoration indicators, such as the peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM). The results are presented in Table 2. The PSNR increased by more than 16.4% with the proposed method compared with the original map, and over 7.5% compared with other methods. At the same time, the SSIM of the proposed method was approximately 0.9, a significant improvement compared with that of other methods.

To evaluate the image sharpness after dehazing, three quantitative indices (grayscale mean gradient (GMG), standard

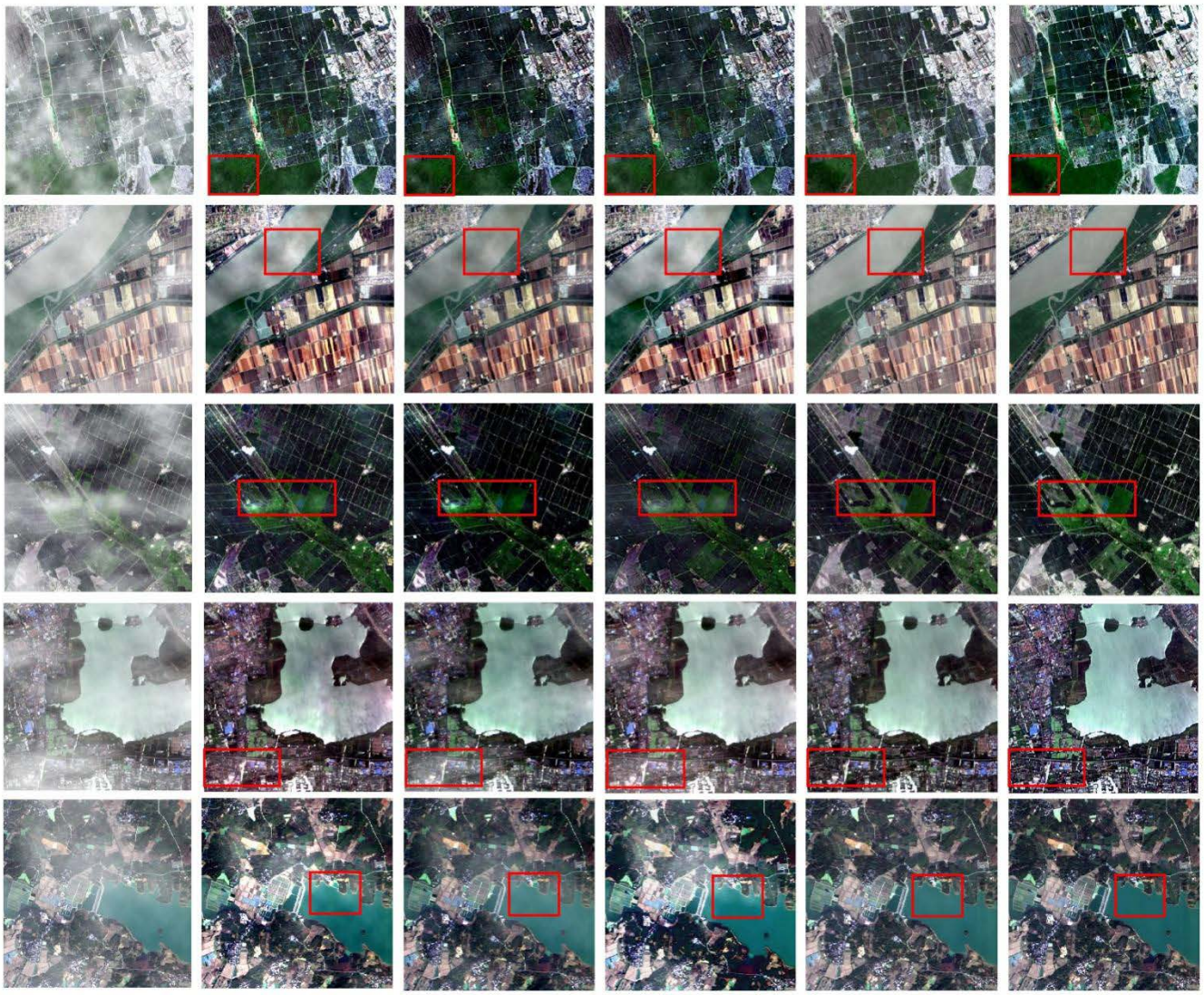

FIGURE 6. Dehazing effects of synthetic hazy images. (a) Synthetic hazy image; (b) Dehazing effect of dark channel method; (c) DehazeNet dehazing effect; (d) AOD-NET dehazing effect; (e) Dehazing effect of the proposed method; (f) Haze-free image.

TABLE 2. Image index for dehazing effect.

Evaluation Index	Hazed image	DCP	Dehaze Net	AOD-NET	Proposed
PSNR	I1:58.05	I1:62.899	I1:61.983	I1:65.025	**I1:72.251**
	I2:60.003	I2:63.708	I2:62.021	I2:65.443	**I2:68.584**
	I3:56.268	I3:65.988	I3:65.942	I3:63.146	**I3:72.238**
	I4:61.849	I4:65.989	I4:64.901	I4:66.181	**I4:70.323**
	I5:60.246	I5:63.836	I5:68.023	I5:62.322	**I5:75.455**
SSIM	I1:0.629	I1:0.655	I1:0.674	I1:0.741	**I1:0.935**
	I2:0.582	I2:0.631	I2:0.687	I2:0.782	**I2:0.893**
	I3:0.546	I3:0.726	I3:0.733	I3:0.712	**I3:0.884**
	I4:0.609	I4:0.802	I4:0.788	I4:0.845	**I4:0.896**
	I5:0.643	I5:0.743	I5:0.812	I5:0.766	**I5:0.923**

deviation (STD), and information entropy (E)) were selected to quantify the experimental results. GMG reflects the contrast of tiny details and texture variations in the images, STD indicates the discrete degree of image pixel gray value relative to the mean value, and E refers to the average amount of information in the images, which measures the amount of information in the images from the perspective of information theory. The higher the E value, the more the information contained in the images. According to Table 3, the proposed method was superior to the other methods for all indicators and significantly improved hazy images. Table 3 shows the images with clearer details and richer colors, processed by the proposed method.

To further verify the proposed method, GF-1 multispectral images with heterogeneous hazes acquired from around Nantong City and Huzhou City in 2013 were selected; the results are shown in Fig. 7. Generally, the experimental results were consistent with the synthetic haze image results. Compared with other methods, the proposed method had lower haze residue level and better recovery effect.

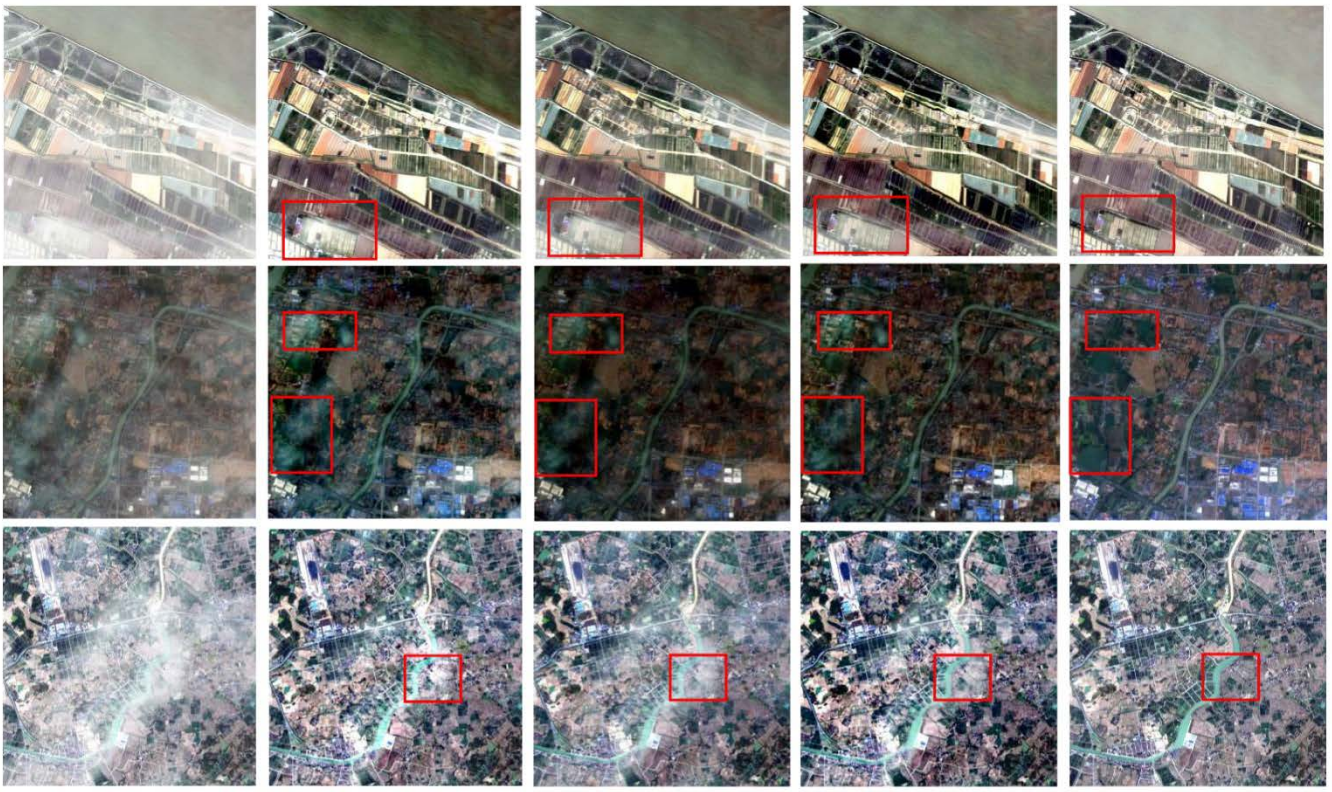

FIGURE 7. Dehazing effects for GF-1 hazy images. (a) Hazy images; (b) Dehazing effect of dark channel method; (c) Dehazing effect of DehazeNet; (d) Dehazing effect of AOD-NET; (e) Dehazing effect of the proposed method.

TABLE 3. Quantitative indexes for the dehazing effect in clarity.

Evaluat-ion index	Hazed image	DCP	Dehaze Net	AOD-NET	Propose d
E	I1:6.574	I1:6.899	I1:6.583	I1: 7.025	I1: **7.176**
	I2:6.203	I2:6.708	I2:6.621	I2: 6.817	I2: **7.084**
	I3:6.868	I3:6.988	I3:6.942	I3: 7.146	I3: **7.238**
	I4:6.849	I4:6.989	I4:6.901	I4: 7.181	I4: **7.323**
	I5:6.246	I5:6.836	I5:7.023	I5:7.322	I5: **7.455**
GMG	I1:6.447	I1:11.343	I1:9.567	I1:10.588	I1:**12.255**
	I2:3.925	**I2:9.021**	I2:7.573	I2: 8.181	I2: 7.753
	I3:5.391	**I3:8.576**	I3:8.029	I3: 7.423	I3: 8.345
	I4:3.605	I4:5.645	I4:5.868	I4: 6.189	I4: **6.304**
	I5:5.833	I5:5.046	I5:4.948	I5: 6.022	I5: **6.045**
STD	I1:29.204	I1:29.441	I1:37.715	I1:39.787	I1: **45.542**
	I2:18.334	I2:28.665	I2:30.308	I2:36.309	I2: **37.304**
	I3:30.369	I3:40.227	I3:38.204	I3:46.070	I3: **47.941**
	I4:33.190	I4:49.689	I4:44.136	I4:**54.656**	I4: 54.194
	I5:36.245	I5:52.276	I5:41.344	I5:57.355	I5: **58.022**

TABLE 4. Impact of different modules on network performance.

Method	Components			PSNR	SSIM
	RB	SA	CA		
Baseline				60.637	0.622
Model-1	✔			65.241	0.726
Model-2	✔	✔		69.364	0.809
Proposed	✔	✔	✔	70.544	0.863

D. ABLATION EXPERIMENT

To further verify the effectiveness of different modules in the network, this study conducted ablation experiments mainly focusing on three key modules in this network: residual block, channel attention, and the impact of spatial attention on network performance. The attention module of the dehazing network was removed and the residual blocks replaced with ordinary convolutional layers as the baseline. Then, each module was added one by one as models 1-2 and the dehazing effect was evaluated, as shown in Table 4. Figure 8 shows the dehazing performance of different ablation models.

The results show that these modules all contributed to the network performance and overall, adding these modules improved PSNR performance by 16.3% compared to the baseline. At the same time, the PSNR performance of the network with residual module and spatial attention module improved by 7.6% and 14.4% compared with the baseline. Also, adding all the modules improved SSIM performance by 0.241 compared to the baseline. The SSIM performance of the network with residual module and spatial attention

FIGURE 8. Dehazing effects for ablation models as described in Table 4 (a) Synthetic hazy image; (b) Dehazing effect of baseline; (c) Dehazing effect of Model-1; (d) Dehazing effect of Model-2; (e) Dehazing effect of the proposed method; (f) Haze-free image method.

module improved by 0.104 and 0.187 compared with the baseline.

IV. CONCLUSION

Conventional remote sensing image dehazing methods require manual selection of features and prior information, resulting in poor stability and universality. Therefore, a trainable end-to-end dehazing neural network for single images was proposed. In the proposed method, the simplified degradation model was skipped, and end-to-end learning was performed directly to fully learn the image features. Simultaneously, the spatial and CA modules were designed for the spatial distribution and channel variation features of hazy remote sensing images based on the attention mechanism. Compared with an ordinary CNN, it showed a powerful capability for feature extraction.

Finally, the heterogeneous haze in some areas of the GF-1 remote sensing images were effectively removed. In both the haze residue level and evaluation index, this method was superior to the conventional dark channel method, DehazeNet, and AOD-NET.

This study has some limitations. Theoretically, the best way to verify the dehazing effect is to compare the images after dehazing with those in the same area on a sunny day; however, in this study, for same area images, the background information in hazy images acquired on different days were different owing to factors such as differences in light conditions. Therefore, this method must be verified using more suitable data that would be obtained from more extensive data sources in future. Quantitative indicators for remote sensing, such as the normalized vegetation index and AOD, were used to evaluate the dehazing effect for remote sensing data and guaranteed the accuracy and reliability of the evaluation results.

ACKNOWLEDGMENT

The authors would like to thank Editage (www.editage.cn) for their English language editing service.

REFERENCES

[1] Q. Qiong, C. Zhang, and Y. Quan, "An adaptive dehazing method for a single remote sensing image considering spatial and spectral differences," *J. Wuhan Univ. Inf. Sci. Ed.*, vol. 44, no. 9, pp. 1369–1376, 2019.

[2] Z. Xu, X. Liu, and N. Ji, "Fog removal from color images using contrast limited adaptive histogram equalization," in *Proc. 2nd Int. Congr. Image Signal Process.*, Tianjin, China, Dec. 2009, pp. 1–5, doi: 10.1109/CISP.2009.5301485.

[3] S. G. Narasimhan and S. K. Nayar, "Contrast restoration of weather degraded images," *IEEE Trans. Pattern Anal. Mach. Learn.*, vol. 25, no. 6, pp. 713–724, Jun. 2003.

[4] J. E. McDonald, "The saturation adjustment in numerical modelling of fog," *J. Atmos. Sci.*, vol. 20, no. 5, pp. 476–478, Sep. 1963.

[5] L. Yu, X. Liu, and G. Liu, "A new dehazing algorithm based on overlapped sub-block homomorphic filtering," *Proc. 8th Int. Conf. Mach. Vision. Int. Soc. Opt. Photon.*, Dec. 2015, Art. no. 987502.

[6] D. J. Jobson, Z.-U. Rahman, and G. A. Woodell, "A multiscale Retinex for bridging the gap between color images and the human observation of scenes," *IEEE Trans. Image Process.*, vol. 6, no. 7, pp. 965–976, Jul. 2002.

[7] Z. Zhu, H. Wei, G. Hu, Y. Li, G. Qi, and N. Mazur, "A novel fast single image dehazing algorithm based on artificial multiexposure image fusion," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–23, 2020.

[8] M. Zheng, G. Qi, Z. Zhu, Y. Li, H. Wei, and Y. Liu, "Image dehazing by an artificial image fusion method based on adaptive structure decomposition," *IEEE Sensors J.*, vol. 20, no. 14, pp. 8062–8072, Jul. 2020.

[9] R. Fattal, "Single image dehazing," *ACM Trans. Graph*, vol. 27, no. 3, pp. 1–9, Aug. 2008.

[10] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341–2353, Dec. 2011.

[11] D. Shubo, X. Wei, and P. Yongjie, "Remote sensing image dehazing method based on dark channel prior," *Acta Optica Sinica*, vol. 37, no. 3, pp. 348–354, 2017.

[12] X. Jiwei, Z. Yang, and X. Xianguang, "A study on dehazing algorithm of DCM-HTM for optical remote sensing images," *J. Geomat. Sci. Technol.*, vol. 33, no. 5, pp. 513–519, 2016.

[13] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "DehazeNet: An end-to-end system for single image haze removal," *IEEE Trans. Image Process.*, vol. 25, no. 11, pp. 5187–5198, Nov. 2016.

[14] W. Ren, S. Liu, and H. Zhang, "Single image dehazing via multi-scale convolutional neural networks," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 154–169.

[15] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "AOD-Net: All-in-one dehazing network," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4770–4778.

[16] 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 *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2019, pp. 1375–1383.

[17] Y. Dong, Y. Liu, and H. Zhang, "FD-GAN: Generative adversarial networks with fusion-discriminator for single image dehazing," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 7, 2020, pp. 10729–10736.

[18] K. Enomoto, K. Sakurada, W. Wang, H. Fukui, M. Matsuoka, R. Nakamura, and N. Kawaguchi, "Filmy cloud removal on satellite imagery with multispectral conditional generative adversarial nets," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 48–56.

[19] H. Jiang and N. Lu, "Multi-scale residual convolutional neural network for haze removal of remote sensing images," *Remote Sens.*, vol. 10, no. 6, p. 945, 2018.

[20] M. Qin, F. Xie, W. Li, Z. Shi, and H. Zhang, "Dehazing for multispectral remote sensing images based on a convolutional neural network with the residual architecture," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 11, no. 5, pp. 1645–1655, May 2018.
[21] A. Mehta, H. Sinha, M. Mandal, and P. Narang, "Domain-aware unsupervised hyperspectral reconstruction for aerial image dehazing," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2021, pp. 413–422.
[22] D. Chao, C. L. Chen, and X. Tang, "Accelerating the super-resolution convolutional neural network," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 391–407.
[23] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Computer Vision*. Cham, Switzerland: Springer, 2016.
[24] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[25] O. Russakovsky, J. Deng, and H. Su, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015.
[26] K. Perlin, "Improving noise," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 681–682, 2002.

ZHIJIE HE was born in 1993. He received the bachelor's degree from Wuhan University, in 2016. He is currently pursuing the Ph.D. degree in physical electronics with the Shanghai Institute of Technical Physics, Chinese Academy of Sciences. His research interests include remote sensing image processing and deep learning.

CAILAN GONG received the Ph.D. degree from the East China Normal University, in 2002. Currently, she works as a Professor with the Key Laboratory of Infrared System Detection and Imaging Technology, Shanghai Institute of Technical Physics, Chinese Academy of Sciences. Her current main research interests include image processing of new remote sensors, the application of machine learning algorithms in remote sensing image processing and information extraction, and the processing and application of UAV hyperspectral remote sensing data. She has participated in several national and Shanghai scientific research projects as the Project Leader and the Main Completer. She has published more than 50 papers, has authorized six invention patents, participated in the preparation of two national standards (released), and participated in the preparation of a monograph *Key Technologies of Remote Sensing Retrieval of Ecological Parameters in Qinghai Lake Basin* (published in 2016). She is a member of the Shanghai Aerospace Remote Sensing Society. She has won the Shanghai Science and Technology Progress Award twice as a main participant.

YONG HU received the bachelor's degree from the Harbin Institute of Technology, in 1999, and the Ph.D. degree from the Shanghai Institute of Technical Physics, Chinese Academy of Sciences, in 2004. Currently, he works as an Associate Professor with the Key Laboratory of Infrared System Detection and Imaging Technology, Shanghai Institute of Technical Physics, Chinese Academy of Sciences. His current research interests include machine vision, deep learning image algorithm, and research and development of domestic satellite data remote sensing cloud computing platforms. As the Project Leader or Technical Backbone, he has completed new satellite data image preprocessing, target recognition, and software development projects, such as Tiangong. He is also the Director of the Shanghai Aerospace Remote Sensing Society and the Head of the Space Science and Application Standardization Working Group, National Space Science and Application Standardization Technical Committee.

LAN LI received the Ph.D. degree from the Shanghai Institute of Technical Physics, Chinese Academy of Sciences, in 2021. Currently, he works as a Postdoctoral Fellow with the Key Laboratory of Infrared System Detection and Imaging Technology, Shanghai Institute of Technical Physics, Chinese Academy of Sciences. His research interests include remote sensing image classification and UAV image processing.