An End-to-End System for Unmanned Aerial Vehicle High-Resolution Remote Sensing Image Haze Removal Algorithm Using Convolution Neural Network

YUFENG LI, JINGBO REN, AND YUFENG HUANG
Department of Electronic Information Engineering, Shenyang Aerospace University, Shenyang 110135, China
Corresponding author: Yufeng Li (liyufeng@sau.edu.cn)

This work was supported by the Project of Liaoning Provincial Department of Education (No.L201701, No.L201735).

ABSTRACT An end-to-end image dehazing method based on convolution neural network is presented to solve the problem in which Unmanned Aerial Vehicle (UAV) high-resolution remote sensing images have reduced image sharpness due to haze. First, the original atmospheric scattering model is adapted to get an end-to-end dehazing model. Then, several unknown parameters are unified into one parameter, and the unknown parameter is estimated by using a multiscale convolution neural network. Finally, the parameter estimates are incorporated into the dehazing model to get a haze-free image. In this article, the haze removal effect for different types of unmanned remote sensing images is tested and compared with those of the main dehazing algorithms. The experiments show that the algorithm in this article has different degrees of improvement regarding its visual effect and objective indicators.

INDEX TERMS End to end image dehazing, deep learning, multiscale convolutional neural network, remote sensing image.

I. INTRODUCTION With the continuous development of Unmanned Aerial Vehicle (UAV) technology, remote sensing images are widely used in military defense, disaster emergency responses and ecological environment monitoring. UAV remote sensing has high scientific value in the refinement of regional information. It has the characteristics of a high spatial resolution, a high frequency, and high cost performance [1]. High-resolution remote sensing image processing has emerged as the times have required. High-resolution remote sensing technology has a great impact on the atmospheric environment, especially on the imaging spectrum in the visible light range. In recent years, due to the existence of haze, images taken in outdoor scenes often suffer from poor visibility, reduced contrast, blurred image quality and color offset [2]. Although processing technology for haze images has been developed for both image enhancement and image restoration, the processing technology for haze images is insufficient. Dehazing technology is a kind of computer vision technology applied in haze weather. It can effectively avoid the degradation of the image quality caused by haze weather and provide better data results for image processing, image analysis and image understanding.

At this stage, the haze removal technology mainly includes three methods, and it has achieved some results:

(1) Image enhancement-based haze removal method: This method uses mature image processing algorithms to enhance haze images in order to uncover the features and achieve the useful value of highlighting the target in the image. The disadvantage of this method is that it will inevitably result in information loss in other parts of the image after highlighting the target features, thus distorting the processed image.

(2) Haze removal method based on a physical model [3]: This method restores image based on the atmospheric scattering model or an improved method of the atmospheric scattering model, solves the inverse process of image degradation...
using a mathematical method, and finally achieves the goal of restoring a clear image [4]. Representative works along this line of research include [5]–[12]. Tan [5] proposes a local contrast maximization dehazing algorithm, which maximizes the local contrast by observing the contrast difference between the clear image and the haze image to achieve the effect of dehazing. Fattal [6] by analyzing the image reflectivity, concludes that the projection image is not locally related to the surface shadow, and through this conclusion, dehazing is realized. He et al. [7] introduces an algorithm based on the dark channel prior, which proposes that there is at least one color channel with the lowest intensity pixels with haze. Berman et al. [8] proposed dividing the color in the image RGB space into clusters and linearly restored the clear image using a prior formula with the help of a line of each color cluster in the RGB space in the haze image. Jiang et al. [9] based on the dark channel prior algorithm, proposed a novel adaptive dual channel prior image dehazing method, which combines the dark channel prior and the bright channel prior. Ju et al. [10] proposed an improved atmospheric scattering model (IASM) in which the transmittance map is directly estimated by linear operations of the brightness, saturation and gradient; and the atmospheric light and scene incident light can be accurately estimated by combining sky related features and the guidance model (GEM). Shu et al. [11] proposed a hybrid regularized variational framework to simultaneously estimate depth map and haze-free image. In particular, introduced the second-order total generalized variation (TGV) regularizer to constrain the estimation of depth map. Zhu et al. [12] proposed a prior algorithm of color attenuation, which is a supervised learning algorithm for image denoising. Liu et al. [13] proposed a unified second-order variational framework to refine the depth map and restore the haze-free image. The second-order framework can preserving important structures in both depth map and haze-free image. Shu et al. [14] proposed a hybrid variational model with promoted regularization terms to refining transmission map, and then using an alternating direction algorithm to obtained final haze-free image.

3) Learning-based methods [15]. In recent years, with the rapid development of depth learning, methods to restore haze images using a convolution neural network based on deep learning have continuously emerged. Representative works along this line of research include [16]–[21]. Tang et al. [16] used a random forest regression to learn the correlation between features and transmittance by collecting multiscale features of images, such as dark channels and the local maximum contrast. Cai et al. [17] proposed the DeHazeNet dehazing network, obtained the weights needed by the network by training, and then estimated the transmission rate of haze images using the forward propagation of the network. Ren et al. [18] used a multiscale neural network, and learned the mapping relationship between a haze image and the transmittance. The results showed that the algorithm has a good effect on synthesizing haze images and real images. Li et al. [19] proposed the AOD-Net network, which clarifies haze images by building an end-to-end model. The idea of this network is to reduce the accumulation error caused by estimating the parameters of the physical model many times. Ren et al. [20] proposed an algorithm that hinges on an end-to-end trainable neural network that consists of an encoder and a decoder. The encoder is exploited to capture the context of the derived input images while the decoder is employed to estimate the contribution of each input to the final dehazed result using the learned representations attributed to the encoder. Liu et al. [21] consists of three modules: preprocessing, the backbone, and postprocessing. The trainable preprocessing module can generate learned inputs with better diversity and more pertinent features. The backbone module implements a novel attention-based multiscale estimation on a grid network. The postprocessing module helps to reduce the artifacts in the final output.

In summary, learning-based methods have gradually become the mainstream image processing methods in recent years, and their effects are significant. This article improves the classical atmospheric scattering model for haze images using actual high-resolution remote sensing images from an unmanned aerial vehicle. Several parameters in the model are unified into an input-related variable $K(x)$, which adaptively adjusts to different input images. Thus, the traditional problem of solving the transmittance and the atmospheric light value of the atmospheric scattering model is transformed into the estimation of the unified variable $K(x)$ and the minimization of the error. At this time, an end-to-end deep learning model combining the coarse-scale and fine-scale is built based on the characteristics of high-resolution images. First, a coarse-scale convolution is performed on haze images, and the low frequency feature information is extracted from the images. Then, the features of the coarse-scale network convolution are linearly connected using the fine-scale convolution. Then, the high-frequency feature information is extracted from the image by using the fine scale network, and the information of the haze image variable $K(x)$ is obtained. Finally, variable $K(x)$ of the final minimization error is obtained via deep learning with different training sets. The haze-free image can be obtained by incorporating the improved atmospheric scattering model. Compared with other algorithms, it can be concluded that the algorithm in this article has better performance in haze removal.

II. DEHAZING ALGORITHM BASED ON AN ATMOSPHERIC SCATTERING MODEL

Based on the causes of light scattering and the combined effect of haze on light scattering, as shown in Figure 1, the Nayar and Narasimhan [22] and McCartney [23] believe that the images taken are mainly affected by two reasons.

1) The reflected light of the target is absorbed and scattered by the suspended particles in the medium during the transmission process, which results in energy attenuation. This usually reduces the brightness of the image and the contrast of the image.
(2) Ambient light such as sunlight and other objects’ reflected light is affected by the particles in the medium to form stray light. Stray light is formed by light scattering, which usually makes the captured image blurry, resulting in the image color not being natural. The captured image consists of two parts: one is the reflected light of the attenuated target caused by atmospheric scattering and absorption, and the other is the atmospheric light caused by atmospheric scattering.

The cause of haze formation in an image is represented using the atmospheric scattering model [23], [24], which can be written as follows:

\[ I(x) = J(x)t(x) + A(1 - t(x)) \]  

Here, \( J(x)t(x) \) is the reflected light of the attenuation target caused by atmospheric scattering and absorption, \( A(1 - t(x)) \) is the atmospheric light caused by atmospheric scattering, \( I(x) \) is the haze image, and \( J(x) \) is the clear haze image. There are two important parameters in the atmospheric scattering model. \( A \) is the global atmospheric light, and \( t(x) \) is the transmission matrix, the latter of which is defined as follows:

\[ t(x) = e^{-\beta d(x)} \]  

where \( \beta \) is the atmospheric scattering coefficient and \( d(x) \) is the distance between the subject and the camera. To obtain the final \( J(x) \), the formula is transformed into the following:

\[ J(x) = \frac{1}{t(x)}I(x) - A \frac{1}{t(x)} + A \]  

Most of the existing algorithms follow the following steps to restore a hazy image into a clear image: (1) estimate the transmission matrix \( t(x) \) using the blurred image \( I(x) \), (2) estimate \( A \) with some empirical methods, and (3) estimate the clear Image \( J(x) \) [25] through formula (3). Estimating the transmission matrix and global atmospheric light values from haze images is a problem of uncertainty. Some literatures have proposed using visual cues to capture the statistical features in haze images to approximate the transmission matrix or global atmospheric light values for the images [8], [26]–[28]. However, the image recovered via approximation still differs visually from the original image, which leads to a larger enhancement space for haze removal using the atmospheric scattering model.

III. END-TO-END DEMISTING ALGORITHM BASED ON DEEP LEARNING

A. DEFICIENCY AND IMPROVEMENT OF THE ATMOSPHERIC SCATTERING MODEL

Through observation, it can be found that the noise caused by haze in an image is actually uneven and is related to the attenuation of the scene caused by the haze and the physical distance to the camera surface. If a uniform atmospheric scattering model is used for dehazing, in this case, all the pixels in the image go through the same parameter solving process, and the haze content of the target image is different, resulting in the partial distortion or insufficient dehazing of the image after dehazing. This shows that the dehazing process needs to be changed according to the input image, and the recovery model must also be suitable. As mentioned above, in the two independent steps of the estimation transfer matrix and atmospheric light, it is impossible to completely simulate the process of inverse dehazing of the atmospheric scattering model, which will inevitably lead to parameter estimation errors. Then, the value after the completed parameter estimation will be incorporated in the atmospheric scattering model, and the error will accumulate and may be mutually amplified. Therefore, there are some limitations to the algorithms that use priors or hypotheses.

To avoid the error caused by the independent estimation of parameters \( A \) and \( t(x) \) and to accumulate the amplification error when formula (3) is used for the calculation, inspired by paper [19], the original atmospheric scattering model is improved. As shown in Figure 2, parameters \( A \) and \( t(x) \) are unified in a variable \( K(x) \), and the parameter is estimated to minimize the error. In this way, formula (3) can be written as follows:

\[ J(x) = K(x)I(x) - K(x) + b \]  

Here, we also use the following:

\[ K(x) = \frac{1}{t(x)}(I(x) - A) + (A - b) \]  

In formula (5), \( b \) is a constant with a deviation, and \( J/I(t(x)) \) and \( A \) are integrated in variable \( K(x) \). Since the value of \( K(x) \) depends on the parameters in \( I(x) \), the parameters in \( K(x) \) change with the input haze image, thus minimizing the error between the output Image \( J(x) \) and the real image. The advantage of this method is that it forms an end-to-end dehazing method. That is, if a haze image is input and \( K(x) \) is obtained through the internal calculation of the algorithm,
the haze-free image can be recovered directly without the need to estimate different parameters and then insert them into the atmospheric scattering model similar to other algorithms. This method unifies several parameters that determine the dehazing effect for a variable parameter related to the input image so that the final dehazing effect is related to the haze concentration in the input image, which makes the model have good robustness.

**B. END TO END DEEP LEARNING NETWORK MODEL DESIGN**

To allow the end-to-end model mentioned in the previous section to estimate variable \( K(x) \) more accurately, the algorithm can be adaptively adjusted according to the different input images. With the rapid development of deep learning, a large number of images can be trained through the establishment of a deep learning network so that the trained network parameters can more flexibly adapt to various input images.

Through the analysis of high-resolution images, we can find that the high-resolution images of UAV aerial photography are rich in detailed information. There are not only large areas (such as farmland, water surface, land, etc.) in the image but also a large number of detailed areas (such as residential areas, road traffic networks, ground with textural characteristics, etc.), which result in the ability to use deep learning to extract feature information. We should not only consider extracting low-frequency feature information but also consider extracting high-frequency feature information. To solve this problem, this article proposes a multiscale convolution neural network based on deep learning. The multiscale neural network consists of two parts: one is the coarse-scale neural network for low-frequency information extraction, and the other is the fine-scale neural network for high-frequency information extraction. Through the use of different scale networks for the feature extraction of an input image, the mapping relationship between variable \( K(x) \) mentioned above and the input image is obtained by using deep learning, thus skipping the previous process of solving multiple parameters in the dehazing process, and dehazing can be realized by one parameter variable.

Figure 3 shows the structure of the multiscale neural network established by the algorithm. First, the rough structure of the \( K(x) \) of each image is obtained through the coarse-scale network, and then it is refined through the fine-scale network. Both the coarse-scale network and the fine-scale network are applied to the original input haze image. In addition, the output of the coarse-scale network is transferred to the fine-scale network as additional information. Therefore, the fine-scale network can refine the details of the coarse-scale network with details.

The network structure of the algorithm of this article mainly includes three parts: convolution layers, pooling layers and upsampling layers.

Convolution layers: The convolution operation is used to extract the local features of the image to filter out the redundant components in the image and get the image features. In the coarse-scale network, it is mainly used to extract the low-frequency feature information of the image. The low-frequency feature of a haze image is mainly a large area or large areas with the same color in the image. Therefore, in the coarse-scale network, three large convolution kernels, \( 11 \times 11, 9 \times 9 \) and \( 7 \times 7 \), are used to expand the receptive field of the network to extract the features of the large-scale image information. The \( 11 \times 11 \) and \( 9 \times 9 \) convolution kernels have five layers, and the \( 7 \times 7 \) convolutional kernel has ten layers. By reducing the number of convolution kernels, the low-frequency feature information of an image is extracted gradually and effectively. The output of the \( 7 \times 7 \) convolution kernel is used as the additional information of the fine-scale network, and more layers can be used to ensure better matching with the fine-scale network when it is connected linearly. In addition, due to the large scale of the convolution kernels, in order to avoid an exploding gradient and slow convergence of the whole network, this article introduces residual learning in the coarse-scale network. The residual network connects the input directly to the later convolution layer through other paths so that the latter convolution layer can carry out residual learning. Through this kind of jump connection, the integrity of the image features can be maintained, and the accuracy of the feature extraction of the coarse-scale network can be improved. The low-frequency feature information is extracted from the image after the coarse-scale network, but for the high-resolution image, there is still a lot of high-frequency information, such as land textures, road traffic networks and so on. At this time, the convolution kernel of the receptive field cannot meet the requirements of extracting the high-frequency feature information, which needs to be completed through the fine-scale network. In the fine-scale network, in order to connect to the feature image from the coarse-scale network linearly, the image is first passed through a \( 7 \times 7 \) convolution kernel, which is the same as that of the coarse-scale network, but there are
only 4 layers here. Then, the output of the four feature images and the coarse-scale network output a total of five feature images connected to 5 layers of $5 \times 5$ convolution kernels and 10 layers of $3 \times 3$ convolution kernels. Decreasing the number of convolution kernels can effectively extract the high-frequency feature information of the image step by step. In this way, after the coarse-scale network and fine-scale network, the feature map not only contains low-frequency feature information but also has high-frequency feature information, which makes it ready for the later deep learning training. The network parameters are shown in Table 1.

The convolution formula of each layer mentioned in this article is as follows:

$$f_{l+1}^n = \sigma \left( \sum_m (w_{m,n}^{l+1} \ast f_m^l) + b_{n}^{l+1} \right)$$

where $f_{l}^n$ and $f_{l+1}^n$ are characteristic graphs of the current layer $L$ and the next layer $L+1$, respectively; $w$ and $b$ represent the convolution kernel and offset, respectively; and the symbol $\ast$ represents convolution operation. To avoid the lack of nonlinear factors, the $\sigma$ activation function is added to enhance the nonlinear expression ability of the model. In this article, the activation function adopts the random rectified linear unit (RReLU). Using the RReLU neuron as the activation function is an improvement over the ReLU neuron and is more flexible than the ReLU neuron [20]. As shown in Figure 4, the slope of the RReLU neurons in the negative area is not fixed and randomly selected in a given range, which will be determined according to the input in the test phase. The advantage is that it can more effectively avoid the occurrence of the “neuron death” problem in the back propagation process. The formula of the RReLU neuron function is as follows:

$$y_{ji} = \begin{cases} x_{ji} & \text{if } x_{ji} \geq 0 \\ a_{ji}x_{ji} & \text{if } x_{ji} < 0 \end{cases}$$

where $a_{ji} \sim U(l, u)$, $l > u$, and $u \in [0,1]$.

### TABLE 1. The network parameters.

| NETWORK         | CONVOLUTION KERNELS | LAYERS |
|-----------------|---------------------|--------|
| COARSE-SCALE NETWORK | 11x11               | 5      |
|                 | 9x9                 | 5      |
|                 | 7x7                 | 10     |
| FINE-SCALE NETWORK | 7x7                 | 4      |
|                 | 5x5                 | 5      |
|                 | 3x3                 | 10     |

### FIGURE 4. Activation function comparison.

### FIGURE 5. End to end UAV high-resolution remote sensing image dehazing algorithm based on convolution neural network.

Pooling layers: Since the structure of a high-resolution remote sensing image is complex and the content of an image has a high number of pixels, the image contains much feature information after the convolution layers. To further reduce the dimension of the feature image and retain the features, a $2 \times 2$ max pooling layer is inserted after each convolution layer.

Upsampling layers: After passing through a large pooling layer, the size of an image feature map is reduced. To ensure that the size of the feature map and the input image is the same, upsampling layers are added after passing through the maximum pooling layers of the upper layer [18].

### C. ALGORITHM STEPS

The end-to-end UAV high-resolution remote sensing image dehazing algorithm based on deep learning proposed in this article is shown in Figure 5. The specific steps are as follows:

1. The data set is used to train the deep learning neural network, and the mapping relationship between the feature graph variable $K(x)$ and the input image is obtained;
2. Input the haze image into the network model, and get the corresponding characteristic graph variable $K(x)$; and
3. The original haze image and the characteristic image variable $K(x)$ are introduced into the improved atmospheric scattering model, and the haze-free image is obtained.

### IV. EXPERIMENTAL RESULTS AND ANALYSIS

#### A. DATA SET DESIGN AND TRAINING IMPLEMENTATION

The diversity of data sets is an important condition to determine the results of network training. At present, there is no unified standardized data set to use. Through the analysis of the target haze pictures, it is found that the biggest difference between the high-resolution pictures taken by a UAV and the training data set used by the existing algorithms is that there is no sky area, which is a requirement for the selection of data sets. As shown in Figure 6 (a), we first select 27256 indoor images from the NYU2 depth database [25] as the main training set of this network, and the images in the NYU2 depth database include indoor haze free images and their combined haze images, which allow us to better train the network. In addition, as shown in Figure 6 (b), 1100 high-resolution haze images taken by a UAV over the Panjin red beach wetland of Liaoning Province in 2018 are also used, and 1000 of them are selected as the original self-built training set of this algorithm. Since all the pictures are original pictures...
with haze, the training flow of the algorithm in this article is used, as shown in Figure 7. First, we use the NYU2 depth database to conduct preliminary training on the network, then we conduct the dehazing process on all the images in the original self-built training set, and finally we synthesize the data sets with different concentrations of haze through formula (1). In this article, the atmospheric light value $a$ is set between [0.6, 1.0], and $\beta$ of the atmospheric scattering coefficient is {0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6}. The synthesized data set is trained for a second time.

During training, the initialization parameters use Gaussian random variables. The momentum of the back propagation of the neural network is set as 0.9 and the attenuation coefficient is set as 0.0001. The initial learning rate is set to 0.001. After training using the NYU2 deep database, the set learning rate is reduced by half so that the network can train using the subsequent self-built training set more effectively. In this article, the mean squared error loss function (MSE) is taken as the cost function. Then, the goal of network optimization is as follows:

$$L(\theta) = \min_{\theta} \frac{1}{n} \sum_{i=1}^{n} \| t - k(x_i; \theta) \|^2$$ (8)

Here, $\theta$ represents the weight and super parameter of the network to be learned, and the back propagation gradient descent method is used to optimize its parameter in the network training process. The training loss function (8) is used for both the coarse-scale network and the fine-scale network.

**B. EXPERIMENTAL COMPARISON**

In this section, we will compare this algorithm with the most classical and advanced algorithms in terms of its visual effect and objective indicators, and the algorithms that need to be trained all use the databases used in this article. The algorithms involved are as follows:

1. Dark Channel Prior (DCP) - He et al. [7] prior haze removal,
2. Multi scale Convolutional Neural Network (MSCNN) - Ren et al. [18] multi scale convolutional neural network for haze removal,
3. Dehazenet - Cai et al. [17] dehazing based on a CNN,
4. All-in-one Dehazing Network (AOD net) - Li et al. [19] single image end-to-end CNN image dehazing,
5. Dense Connected Pyramid Dehazing Network (DCPDN) - Zhang and Patel [29] dense linked pyramid dehazing,
6. Gated Fusion Network(GFN) - Ren et al. [20] Gated fusion network for single image dehazing, and
7. GridDehazeNet – Liu et al. [21] Griddehazenet: Attention-based multiscale network for image dehazing.

The algorithm in this article uses a deep learning framework named Pytorch, and the training and testing of the network are completed in the Pytorch environment. The hardware environment is an Intel Core i7-8750 h CPU and an NVIDIA GeForce GTX 1050Ti graphics card.

1) COMPARISON OF THE VISUAL EFFECTS ON SYNTHETIC DATASETS

This article first conducts testing using the synthetic dataset and compares the results with those of the seven algorithms mentioned above. The DCP is a prior-based method while the
other methods are based on data training. For a fair comparison, the above data-based training data are the same as the training data used in this article. In this article, both the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) are used to quantitatively evaluate the dehazed images, which are explained in detail in the chapter comparing the objective indicators. From the image shown in Figure 8 (b), it can be seen that the result of the DCP is darker than that of the real image and the other dehazed images, and the DCP can cause severe color distortion. For the MSCNN and DehazeNet in Figures 8 (c)-(d), respectively, the visible haze is still not effectively removed, and the output image color is whiter due to the residual haze. The AOD-Net and DCPND in Figures 8 (e)-(f), respectively, overcome the effect of color distortion to a large extent. The output image color is closer to the real object, but the algorithm can easily generate halos and artifacts around the object, and the blurred parts in the original image are not significantly removed. The GFN and GridDehazeNet are the haze removal algorithms proposed in the last two years. The advantages of these two algorithms are that they successfully suppress artifacts and halos to a certain extent, are visually closer to haze-free images and have better haze removal effects for scenes close to the lens. However, haze can still be clearly seen in scenes with deeper depths, that is, scenes far from the lens. Compared with the existing technology, this article’s proposed method largely compensates for the disadvantage of different degrees of distant and near-range haze removal. This is because the convolution neural network combined with the coarse and fine scales selected in this article can learn the haze images of different scales separately. The coarse scale network is responsible for the lower-resolution part of the image, and the fine-scale network is responsible for the higher-resolution part of the image.

During the recovery process, the images with different depths of field are more specifically dehazed, which shows better robustness and a better visual effect.

2) QUALITATIVE COMPARISONS ON REAL-WORLD IMAGES
To further evaluate the proposed method and to compare the similarities and differences between our algorithm and the other algorithms on unmanned remote sensing image fog removal, the unmanned remote sensing image of Panjin Red Beach dataset in Liaoning Province was used as the test image. Figure 9 shows a comparison of the algorithms. On the synthetic datasets, the DCP can cause color distortion, which is particularly evident in the first row of images in Figure 9 (b). This is mainly because the DCP dehazing algorithm is based on prior knowledge, but the inherent disadvantage is that it relies too much on prior knowledge and ignores the information of the image itself, which results in the dehazing effect being greatly reduced. A similar problem in the MSCNN, DehazeNet and DCPND is that the image after haze removal still has visible haze, which is more evident in the MSCNN, especially in the first row of Figure 9 (c); furthermore, the MSCNN also has color distortion. As we can see from the last three rows of the image in Figure 9 (c), the image after haze removal is significantly different from the original image. The result of the DCPDN algorithm is shown in Figure 9 (f). The white area is overexposed, which makes the color of the building area brighter. The AOD-Net method has the disadvantage of overenhancement, that is, the image itself has a good haze removal effect and the details of haze removal are clear, but the whole image is black. The image in Figure 9 (e) is visually much darker than those of the other algorithms. The GFN and GridDehazeNet algorithms show prominent visual effects after haze removal, which
conforms to people’s intuitive perception of haze-free images. However, for the high-resolution remote sensing images from an unmanned aerial vehicle, the biggest feature is the large resolution, especially in farmland and beach areas. The image has rich textural details, more edge information, and a large color change in a single image. After enlarging the dehazed image with the GFN and GridDehazeNet algorithms, you can see that many textural details are blurred, many areas are blurred, and the color is darker. As a result, the dark areas in the original image are much darker, such as the fourth line images in Figures 9 (g) - (f), which appear black after dehazing. In comparison, the algorithm can clearly see the textural details of the red area on the right half through the fourth line of the picture in Figure 9 (i). Compared with the above methods, our method has a clear overall haze removal effect, retains the maximum amount of scene details, and is more effective at fog removal.

To further evaluate the proposed method, we also use part of the real image data set from Ren et al. [18] and Fattal [30] to compare the seven algorithms mentioned above with those in this article. Figure 10 (a) is a true image with haze. Same as the Red Beach dataset, the DCP algorithm exhibits varying degrees of color distortion, which is evident in the third row of Figure (10) (b). The MSCNN, DehazeNet and DCPND have the same problem as with the Red Beach dataset, that is, there is still some haze in the images. The AOD-Net algorithm has a darker image than other images, but it generally has a good haze removal effect. The GFN and GridDehazeNet algorithms have little difference in their haze removal effects, but there is some residual haze on
the edges of the image. By contrast, our method has clear haze removal. The image details are enhanced appropriately after haze removal to maximize the restoration of the scene’s textural details.

3) OBJECTIVE INDEX COMPARISON
Table 2 shows the average PSNR and SSIM values of dehazed results on the synthetic dataset. It can be seen that the method proposed in this article is much better than the comparison algorithms listed in the paper.

The pictures taken by the UAV are all haze images without clear haze images. The SSIM and PNSR are commonly used to evaluate images that have original clear images. However, in order to compare the dehazing effect of the algorithms more objectively, the MG (mean gradient), ES (edge strength), IE (information entropy) and VAR (variance) are selected in this article and used to analyze the images quantitatively. Table 3 shows the comparison of the objective indicators. The mean gradient of an image refers to the change rate of the gray level of an image, which reflects the change rate of the contrast of the small details of an image. The edge strength can be used to express the sharpness of an image. The larger the value is, the richer the details are, and the clearer the texture. The edge strength represents the magnitude of the gradient of an image’s edge points, which is similar to the average gradient. The larger the value is, the clearer the details are expressed. The image information entropy is the value that represents the whole amount of image information. The larger the value is, the greater the amount of image information quantity, the better the image quality, and the richer the texture. The variance is used to reflect the image color and contrast. The larger the value is, the more prominent the color performance of the image. Through the above four indexes, the above 7 UAV images are analyzed to get the average values. Table 2 shows the parameter indexes of each algorithm after dehazing. It can be seen that the algorithm in this article improves the performance with respect to the mean gradient, edge strength, information entropy and variance data, which shows that the image processed by the algorithm in this article has higher definition, rich detail information, and higher color saturation; therefore, a good dehazing effect is obtained.

In addition, as shown in Table 4, this article also measures the code running time for the 7 test images mentioned above and takes the weighted average value as the average time of a single image dehazing process. It can be seen that the algorithm in this article has both a good processing effect and shorter running time than the other algorithms, ensuring the image dehazing efficiency.

V. ANALYSIS AND DISCUSSIONS
A. EFFECTIVENESS OF THE MULTISCALE NEURAL NETWORK
In this section, we analyze how the multiscale neural network proposed in this article can make the image dehazing better. The combination of the coarse-scale network and fine-scale network can effectively improve the image dehazing effect, and the output of the coarse-scale network can provide additional information for the fine-scale network. Figures 10 (b) and (c) show the images containing only the coarse-scale network and the multiscale network constructed in this article. Figures 11 (d) and (f) show the dehazed images containing only the coarse-scale network and the multiscale network proposed in this article. It can be seen that for the dehazed image that only uses the coarse-scale network, the details are fuzzy, and only the details over large areas can be obtained. In addition, the image effect is not obvious after the final dehazing, and the textural details are still obstructed by a large amount of haze. However, the multiscale network constructed in this article retains the details and
edge information in detail processing, which also proves that only using the coarse-scale neural network cannot effectively extract the details in the image. Compared with the coarse scale network, the multiscale neural network can solve this problem effectively.

**B. EFFECTIVENESS OF THE UNIFIED VARIABLE K(x)**

To illustrate the effectiveness of the K(x) module used in this article, we will compare it with the MSCNN algorithm. This is because the MSCNN algorithm only uses the method based on the convolution neural network. By learning the mapping relationship between the blurred image and its corresponding transmission image, image dehazing is achieved. In this article, the K(x) and multiscale neural network algorithm is used. Through the comparison of Figures 12 (b) and (c), we can see that although the MSCNN algorithm performs well at dehazing, the color distortion phenomenon appears, and the image edges appear hazy, which indicate that the edge processing effect of the algorithm is not very good. The algorithm in this article includes parameter K(x) of the atmospheric scattering model. The results show that the image color is consistent with the original image, and the overall image dehazing effect is consistent, which shows that the atmospheric scattering model K(x) can effectively suppress color distortion, and can improve the overall dehazing effect of the image. Regarding the image edges, it can still achieve a good dehazing effect.

## VI. CONCLUSION

This article proposes end-to-end dehazing technology based on deep learning, which is inspired by the extensive application of deep learning in the field of image processing. Using the multiscale convolution neural network, it can remove haze more effectively, and the introduction of residual learning can reduce the computational load and speed up learning. This article analyzes the differences between the existing dehazing algorithms and the algorithm in this article by comparing their visual effects and objective indexes. The experiment shows that the algorithm in this article has a good effect in all aspects and can meet the requirements of UAV high-resolution remote sensing image dehazing. However, for an image with an uneven haze distribution, the processing speed and processing effect can be further improved. At present, the requirement of real-time processing for UAV image is gradually improved. Our future work will focus on improving the real-time performance of UAV image dehazing, accelerating the processing speed of the algorithm and improving the processing efficiency.

### REFERENCES

[1] X. Liao, Q. Xiao, and B. Zhang, “Unmanned aerial vehicle remote sensing: The trend of popularization and expansion of application,” *J. Remote Sens.*, vol. 23, no. 6, pp. 1046–1052, 2019.

[2] J. Xiao, M. Shen, and J. Lei, “Image conversion algorithm based on generating haze scene against network,” *J. Comput. Sci.*, vol. 43, no. 1, pp. 1–11, Nov. 2019.

[3] X.-J. Huang, Y.-D. Lai, and F. Chen, “Fast single image haze removal algorithm,” *J. Comput. Appl.*, vol. 30, no. 11, pp. 3028–3031, Dec. 2010.

[4] M. Du and H. Sun, “Removal algorithm of uneven haze in unmanned aerial vehicle remote sensing image,” *Surv. Spatial Geographic Inf.*, vol. 42, no. 1, pp. 174–176, 2019.

[5] R. T. Tan, “Visibility in bad weather from a single image,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[6] R. Fattal, “Single image dehazing,” *ACM Trans. Graph. (TOG)*, vol. 27, no. 3, pp. 72, 2008.

[7] 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, Jun. 2009.

[8] D. Berman, T. Treibitz, and S. Avidan, “Non-local image dehazing,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1674–1682.

[9] Y. Jiang, C. Sun, Y. Zhao, and L. Yang, “Image dehazing using adaptive bichannel priors on superpixels,” *Comput. Vis. Image Understand.*, vol. 165, pp. 17–32, Dec. 2017.

[10] M. Ju, Z. Gu, and D. Zhang, “Single image haze removal based on the improved atmospheric scattering model,” *Neurocomputing*, vol. 260, pp. 180–191, Oct. 2017.

[11] Q. Shu, C. Wu, Q. Zhong, and R. W. Liu, “Alternating minimization algorithm for hybrid regularized variational image dehazing,” *Optik*, vol. 185, pp. 943–956, May 2019.

[12] Q. Zhu, J. Mai, and L. Shao, “A fast single image haze removal algorithm using color attenuation prior,” *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3522–3533, Nov. 2015.

[13] R. W. Liu, S. Xiong, and H. Wu, “A second-order variational framework for joint depth map estimation and image dehazing,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 1433–1437.

[14] Q. Shu, C. Wu, Z. Xiao, and R. W. Liu, “Variational regularized transmission refinement for image dehazing,” in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 2781–2785.

[15] R. Li, J. Pan, Z. Li, and J. Tang, “Single image dehazing via conditional generative adversarial network,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8202–8211.

[16] K. Tang, J. Yang, and J. Wang, “Investigating haze-relevant features in a learning framework for image dehazing,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2995–3000.

[17] B. Cui, 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.

[18] G. Tang, L. Zhao, R. Jiang, and X. Zhang, “Single image dehazing via lightweight multi-scale networks,” in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2019, pp. 5062–5069.

[19] 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.

[20] W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W. Liu, and M.-H. Yang, “Gated fusion network for single image dehazing,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 2523–2531.

[21] X. Liu, Y. Ma, Z. Shi, and J. Chen, “GridDehazeNet: Attention-based multi-scale network for image dehazing,” in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 7314–7323.

[22] S. K. Nayar and S. G. Narasimhan, “Vision in bad weather,” in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, Sep. 1999, pp. 820–827.

[23] E. J. McCartney, *Optics of the Atmosphere: Scattering by Molecules and Particles*, vol. 421. New York, NY, USA: Wiley, 1976.

[24] S. G. Narasimhan and S. K. Nayar, “Chromatic framework for vision in bad weather,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. CVPR*, Jun. 2000, pp. 598–605.

[25] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang, “Benchmarking single-image dehazing and beyond,” *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 492–505, Jan. 2019.

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

**FIGURE 12.** Effectiveness of the unified variable k(x). (a) is original image, (b) is dehazed image of the MSCNN algorithm, and (c) is the dehazed image of our algorithm.
[27] N. Hautiere, J.-P. Tarel, and D. Aubert, “Towards fog-free in-vehicle vision systems through contrast restoration,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2007, pp. 1–8.

[28] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, “Instant dehazing of images using polarization,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. CVPR, Dec. 2001, p. 1.

[29] H. Zhang and V. M. Patel, “Densely connected pyramid dehazing network,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2018, pp. 3194–3203.

[30] R. Fattal, “Dehazing using color-lines,” ACM Trans. Graph., vol. 34, no. 1, pp. 13:1–13:14, Dec. 2014.

YUFENG LI received the Ph.D. degree. He graduated in communication engineering from Northern Traffic University, in July 1991, and in communication and information system from Jilin University, in March 2002.

In September 2009, he was a Visiting Scholar with the Beijing University of Technology. In May 2010, he entered the Postdoctoral Mobile Station of the State Key Laboratory of Mobile Communication, Southeast University. He is currently a Professor and a Master’s Supervisor with the School of Electronic Information Engineering, Shenyang University of Aeronautics and Astronautics.

YINGBO REN graduated from the Nanchang University of Aeronautics and Astronautics, in 2017. He is currently pursuing the degree with the Shenyang University of Aeronautics and Astronautics.

YUFENG HUANG received the bachelor’s degree from the Hefei University of Technology and the master’s degree from the China University of Science and Technology.

She is currently a Teacher with the Shenyang University of Aeronautics and Astronautics.