Journal of Computer Science and Electrical Engineering
ISSN: 2663-1946
DOI: https://doi.org/10.61784/jcsee3055

SINGLE IMAGE DEHAZING BASED ON IMPROVED AOD-NET

JinBo Yi
College of Information Technology, Shanghai Ocean University, Shanghai 201306, China.
Corresponding Email: y330456728@163.com

Abstract: Hazy images often suffer from low quality when processed by traditional defogging algorithms, which fail to effectively remove haze. To address this issue, this article propose a novel single-image dehazing model based on the AOD-Net architecture. The model leverages depthwise separable convolutions to construct a lightweight deep neural network. Additionally, this article introduces a new conditional convolution and attention mechanism to enhance feature extraction, thereby improving the network's ability to capture global information from hazy images. To optimize model performance, this article train the proposed model on the NYU dataset and conduct extensive experiments on the same dataset. The dehazing effectiveness is evaluated using full-reference image quality assessment metrics. Experimental results demonstrate that the improved model achieves higher accuracy in dehazing quality compared to existing methods. Furthermore, the incorporation of the new feature extraction module and attention mechanism significantly enhances performance in haze removal, color restoration, and detail preservation, outperforming the original AOD-Net and other traditional approaches.The application feasibility of this technique is extensive: In the domain of autonomous driving, it can enhance the target detection accuracy of on-board cameras in foggy weather; in remote sensing monitoring, it facilitates satellites and unmanned aerial vehicles to obtain clearer surface information; in the area of security surveillance, it can strengthen the reliability of video analysis in low-visibility circumstances. Additionally, the lightweight design of the model can be adapted to edge computing devices, providing technical support for real-time defogging and possessing significant engineering application value and commercial potential.
Keywords: Deep learning; Image dehazing; Atmospheric scattering model

1 INTRODUCTION

Hazy conditions, caused by air pollution, dust, smoke, and atmospheric particles, introduce a complex form of non-linear noise that significantly degrades image quality[1]. This degradation results from the scattering and absorption of light by suspended particles in the atmosphere, leading to reduced contrast, color distortion, and obscured details. The presence of haze poses considerable challenges for various computer vision tasks, including object detection, recognition, and segmentation, all of which rely on high-quality input data for accurate and reliable performance[2-3]. Consequently, the development of effective image dehazing techniques is crucial for enhancing image clarity and improving the robustness of vision-based applications.

Image dehazing has gained widespread importance in diverse fields such as autonomous driving, medical imaging, and video surveillance. In autonomous driving, the ability to accurately perceive lane markings, traffic signs, and pedestrians under adverse weather conditions is essential for ensuring vehicle safety and reliability. Similarly, in video surveillance systems, dehazing algorithms play a vital role in maintaining the clarity of recorded footage, enabling better monitoring and security analysis in low-visibility environments.

Advancements in image processing and deep learning have spurred extensive research into robust dehazing methods capable of addressing the visual distortions introduced by haze. The primary challenges associated with hazy images include reduced contrast, blurred edges, and a loss of critical textural and color information. These issues arise due to the unpredictable scattering of light, which mixes reflected object light with atmospheric particles, causing shifts in perceived color and loss of fine details. Such distortions not only hinder human interpretation of images but also degrade the performance of automated vision systems.

To mitigate these challenges, numerous image dehazing algorithms have been developed to restore image clarity and preserve essential details. Traditional model-based approaches, such as atmospheric scattering models, estimate haze transmission and restore images based on handcrafted priors[4-5]. However, these methods often struggle with complex real-world scenes and may produce artifacts when the assumed priors do not hold. More recently, deep learning-based methods have demonstrated superior performance by leveraging large-scale datasets to learn complex mappings between hazy and haze-free images. Among these approaches, AOD-Net has emerged as a prominent end-to-end deep learning model that directly estimates a transformation function to reconstruct haze-free images[6]. Despite its effectiveness, AOD-Net has limitations in terms of generalization across diverse environments and restoration of fine details.

To address these limitations, this article proposes IAOD-Net, an enhanced deep learning model for single-image dehazing. Building upon AOD-Net, IAOD-Net incorporates several key improvements aimed at enhancing its dehazing capability. First, the article replaces standard convolutional layers with conditional convolution layers, allowing the network to adaptively adjust its filtering behavior based on contextual information. This enables the model to more effectively capture spatially variant haze distributions. Additionally, IAOD-Net integrates two specialized modules: an edge enhancement module and an attention excitation module. The edge enhancement module is designed to recover sharp boundaries and fine textures that are often lost in hazy images, preserving the structural integrity of restored images. Meanwhile, the attention excitation module enhances the network's ability to focus on highly degraded regions, selectively prioritizing areas most affected by haze. This mechanism improves the overall accuracy of haze removal, yielding clearer and more visually natural outputs.

Through extensive experiments, this article demonstrates that IAOD-Net achieves superior performance compared to the original AOD-Net and other state-of-the-art dehazing methods[6]. By integrating adaptive filtering and attention mechanisms, the article approachs effectively restores image details, enhances contrast, and improves color fidelity, making it well-suited for practical applications in autonomous driving, surveillance, and remote sensing.

2 RELATED WORK

Traditional algorithms for image dehazing encompass methods rooted in the Dark null hypothesis (e.g.These techniques, including the Dark Channel Prior and various approaches involving deep learning, frequently encounter challenges in complex, real-world environments and can generate artifacts if the expected assumptions are violated.g.DCPNet, DehazeNet jne).At present, there are primarily three types of algorithms for removing haze from images. The initial group consists of techniques that focus on enhancing the quality of images. This approach overlooks the imaging mechanisms that lead to the degraded images, reinterpreting the haze issue as one of enhancing contrast, thereby accentuating image details and improving the overall contrast. Another approach involves image dehazing algorithms that utilize a physical modeling framework. This method investigates how hazy images are created, constructs a model of the imaging process, and subsequently performs reverse calculations based on this model to retrieve the original clear image. Recently, advancements in deep learning methodologies have found extensive applications in various aspects of image manipulation, including but not limited to image classification, object recognition, and facial recognition. The third type involves most current algorithms for image dehazing that leverage deep learning; they utilize a neural network model to assess the transmittance of a hazy image, subsequently determine the atmospheric light value independently, and finally generate a haze-free image by applying the atmospheric scattering model. Nevertheless, the accuracy of such assessments can sometimes be questionable.

In the process of DCP, the estimation of atmospheric light relies on established insights derived from the concept of the dark channel[1]. To start, candidate points are identified by selecting the highest-ranked pixels from the ordered dark channel map. These candidate points correspond to positions within the original image, from which the brightness values are extracted. The maximum value from this group of brightness values is then determined to represent the atmospheric light value. In both DehazeNet and MSCNN, an estimation of the transmission map is conducted via the convolutional network framework, while the atmospheric light value derives from established knowledge related to the dark channel. Nonetheless, if an image features objects that closely resemble the atmospheric illumination—such as numerous white items or other bright light sources—the calculated atmospheric brightness could become skewed, leading to a dehazed image that appears overly bright. Moreover, the separate estimation of two crucial elements, namely the transmittance matrix and the atmospheric lighting, might exacerbate the inaccuracies when these components are utilized together. Li et al. An advanced convolutional neural network (CNN) model specifically aimed at dehazing, referred to as the Fully Functional Dehazing Network (AOD-Net)[6], has been proposed to enhance efficiency in this area. The architecture of AOD-Net is founded on a restructured model of atmospheric scattering. Instead of generating a transmittance map, the model produces a clear image. Results indicate that AOD-Net surpasses numerous leading techniques in terms of effectiveness.

3 ENHANCED DESIGN OF THE IAOD-NET MODEL

3.1 Equation for Atmospheric Scattering Deformation Model

To elucidate the generation of haze images, the atmospheric scattering model was initially introduced by McCartney[7], and it received further enhancements from Narasinghan and Nayar[8,9]. The formal representation of the atmospheric scattering model can be articulated as follows.

$$\mathrm{I(X)} = \mathrm{J(X)t(X)} + \mathrm{A(1-t(X))} \quad (1)$$

Where $\mathrm{I(X)}$ is the observed intensity, J(x) is the light intensity from the scene object before scattering, $\mathrm{t(X)}$ is the scene transmittance, which represents the amount of light that reaches the observer after scattering, and A represents the global ambient illumination. In addition, $\mathrm{t(X)}$ is the intermediate transmission matrix, defined as

$$\mathrm{t(X)} = e^{-\beta(\gamma)d(X)} \quad (2)$$

Where β is the atmospheric scattering coefficient, and the uniform concentration of haze can be approximated as a constant; c is the wavelength of the reflected light; $\mathrm{d(X)}$ is the scene depth, which is the distance between the corresponding object in the scene and the imaging device.

$\mathrm{K(X)}$To recuperate a clear image from a hazy one, it is essential to determine the light transmission values present in the hazy image and the associated atmospheric illumination from the perspective of the atmospheric scattering framework. Li et al. introduced a new variable by deformating the atmospheric scattering model, so that the neural network model could directly estimate the joint value of transmittance and atmospheric light[10]. Utilizing the atmospheric scattering framework (1), one can express the clean image produced by the proposed network as follows:

$$\mathrm{J(X)} = \frac{1}{\mathrm{t(X)}}\mathrm{I(X)} - \mathrm{A}\frac{1}{\mathrm{t(X)}} + \mathrm{A} \quad (3)$$

$$J(X) = K(X)I(X) - K(X) + b \tag{4}$$

$$K(X) = \frac{(1/t(X))(I(X)-A) + (A-b)}{I(X)-1} \tag{5}$$

Estimating atmospheric light and transmittance together helps to mitigate the issue of overestimating atmospheric light caused by bright regions or sky sections.

3.2 Challenges Associated with the Initial AOD-Net

The AOD-Net architecture struggles with effectively retrieving fine details in images. The network architecture's failure to adapt to varying lighting conditions results in inconsistent efficacy across different illumination scenarios[11-12]. Estimating the atmospheric illumination and transmittance is fundamental to the process of dehazing rooted in the atmospheric scattering theory. Nevertheless, determining the values for transmittance and atmospheric illumination with precision remains a challenging endeavor. In DCP, the atmospheric light level is derived using knowledge from the dark channel prior methodology. If the image features numerous bright elements or additional light sources, the derived atmospheric light value may be excessively high, leading to overexposure in the resulting dehazed image[15-17]. This observation was confirmed through experimentation. It can be observed that utilizing a comprehensive model allows for the recovery of an image free from haze, eliminating the need to individually assess the values of transmittance and atmospheric illumination[18]. Moreover, removal of haze from images plays a vital role in sophisticated computer vision applications including the recognition of obscured objects and the processing of hazy videos. Consequently, there are increasing expectations for the dehazing model regarding its computational efficiency and overall size.

3.3 Refinement of the improvement section

3.3.1 Network structure optimization

In pursuit of this goal, an additional conditional convolution is integrated into the five-layer convolution, drawing upon the transformation equation of the atmospheric scattering model. The architecture can be divided into two main components: an estimation module for K that derives K(x) from the input I(x), and a subsequent module for generating a clean image that leverages K(x) as an adaptive parameter for estimating J(x).

The core element of the innovative model is the K estimation module, which plays a crucial role in determining both the depth and the extent of haze present. Figure 1 illustrates how the article incorporated conditional convolution into the existing five convolution layers, enabling dynamic adjustments of the convolution kernel to cater to various data types. Additionally, the research introduced a couple of supporting components aimed at enhancing model precision and overall effectiveness. Subsequent to the K estimation component, the output generated from the initial segment is integrated into the deformation equation of the atmospheric scattering model, facilitating the computation required to produce the restored image as outlined in (5). Figure 1 illustrates the design of the updated model.

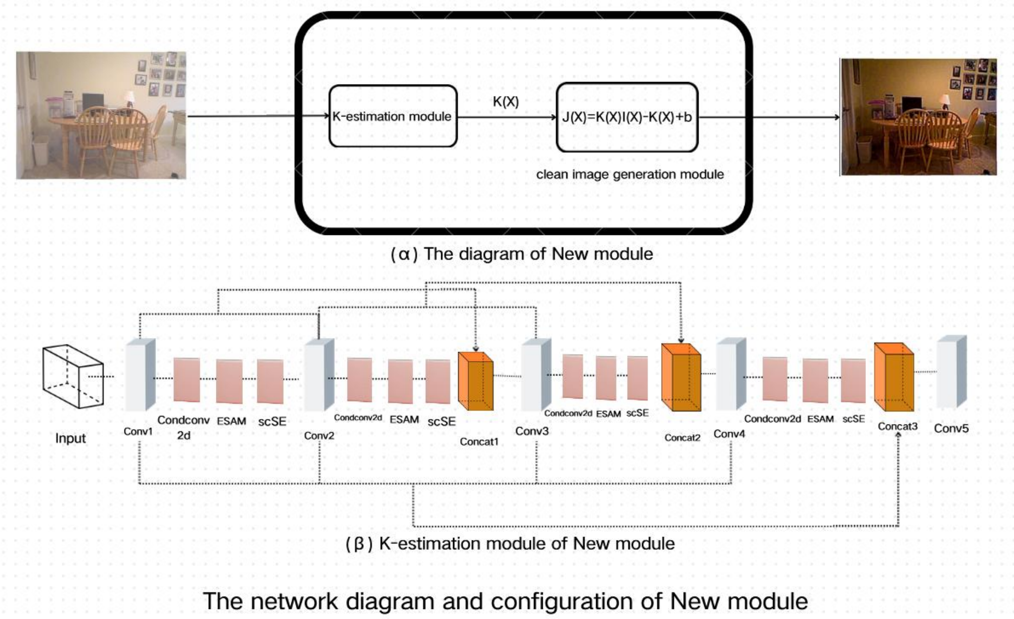

Figure 1 Network Structure Diagram of the New Model

The following provides an in-depth look at how each component is incorporated within the AOD-Net architecture and their respective functionalities: The CondConv2D class establishes a conditional convolutional layer that derives from the _ConvNd base class found in PyTorch. This class employs the _routing helper class to determine the routing weights for every individual input sample, executing the convolution operation by utilizing the weighted sum derived from the routing weights and the corresponding expert kernel.This allows the model to adaptively modify the convolutional filters in response to the input variations, thereby enhancing its capability to recognize diverse patterns within the data, capturing both spatial and channel features with greater flexibility, which in turn boosts the overall efficacy of feature extraction.The Enhanced Squeeze and Excitation Module (ESAM) introduces an edge enhancement mechanism that processes the feature map via convolutional layers and gradient enhancement techniques, thereby augmenting the edge details of the image and aiding in the model's accuracy. By incorporating operations like global average pooling along with convolution, ESAM effectively captures the interrelations among various channels. By flexibly modifying the significance of different features, the model's ability to adapt to intricate scenarios is enhanced.SCSE (Spatial and Channel Squeeze and Excitation) integrates the processes of squeezing and excitation for both channel and spatial aspects, thereby boosting the model's efficiency through the incorporation of attention mechanisms that operate on the input tensor and select the maximum value from their outputs. By acknowledging the interdependence of the spatial and channel dimensions simultaneously, SCSE produces a more intricate set of attention weights. This approach enables the model to recognize intricate interactions among features and enhances its overall effectiveness.

3.3.2 Enhancements in the training methodology

The optimizations in training are as follows: Initially, there is an implementation of a mechanism for adjusting the learning rate dynamically; this mechanism reduces the learning rate over time, based on the performance observed in the validation set. Additionally, the utilization of mixed-precision training through FP16 enhances training efficiency and minimizes memory usage. Third, the approach to storing and retrieving models is refined: This article enhances the processes for saving and loading models during training tasks to prevent memory leaks on the GPU.

4 EXPERIMENTS

This part of the study focuses on assessing the performance of the suggested new model through both qualitative and quantitative methods. The quantitative evaluation takes into account the comprehensive reference image quality assessment indicators, namely PSNR and SSIM.

4.1 Training Data and Experimental Setup

Obtaining blurred images alongside their clear counterparts in natural settings presents a significant challenge. To tackle this, the study creates artificial hazy images using several techniques: It leverages real images that include depth information sourced from the NYU2 indoor depth dataset. Furthermore, the approach involves selecting various atmospheric light parameters A by evenly sampling each color channel within the interval [0.6, 1.0] and choosing $\beta \in \{0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6\}$. For the NYU2 database, Select Article 27,and 256 images are designated as the training set while 3,an additional 170 images were designated as the distinct test set. Moreover, the study also conducted evaluations on images affected by natural haze to assess how well the model generalizes.

The framework was executed utilizing Pytorch version 0.12.1, incorporating five layers of depthwise separable convolutions, four layers utilizing ReLU activation functions positioned subsequent to the convolutional layers, along with three layers that concatenate the data. The experiment utilized an NVIDIA GeForce GTX 1660 Ti 14GB GPU coupled with CUDA version 11.3. During the training phase, the initialization of the weights was performed using random variables drawn from a Gaussian distribution. The framework incorporates ReLU activation functions. Momentum and decay coefficients were established at 0.9 and 0.0001, respectively. with a learning rate fixed at 0.0001 and a batch size consisting of 64 images per batch (480×640). The methodology incorporated a straightforward mean square error (MSE) loss function, which yielded enhancements not only in the Peak Signal-to-Noise Ratio (PSNR), but also enhanced the index of structural similarity (SSIM) along with the overall visual quality.

The recent model requires approximately 860 cycles of training to achieve a suitable fit, and typically shows satisfactory performance after this number of iterations, with peak accuracy observed at around 870 cycles. In this study, the model underwent training for a total of 1000 epochs. Furthermore, it was discovered that constraining the gradient within the interval of $[-0.1, 0.1]$ proved beneficial in the training process. This approach is commonly utilized to ensure stability during the training of recurrent networks.

4.2 Generated Outcomes from the Designated Evaluation Dataset

In order to assess the proposed model's efficiency, experiments were performed utilizing the nyu2 dataset, demonstrating the method’s capabilities in contrast to other leading techniques. The researchers employed synthetic hazy images as inputs for the new model and assessed its results against those from actual images. The test of the proposed technique was conducted utilizing the NYU2 indoor dataset. The evaluation metrics utilized for comprehensive analysis include both PSNR and SSIM. By closely examining the specifics, it becomes evident from figure (b) that the performance of the new model aligns more closely with human visual perception.

The preceding sections analyze the proposed model against the DCP utilizing no-reference metrics for image quality assessment on synthetic datasets. This section utilizes full-reference image quality assessment metrics such as PSNR and SSIM to analyze the performance of the models with respect to the images. Table 1 presents the average values for PSNR and SSIM. Table 1 illustrates that in the synthetic nyu2 test set, the PSNR and SSIM metrics of the new model marginally surpass those recorded for DCP[1].

The study performs tests to evaluate the enhanced model in relation to both the conventional Dark Channel Prior technique and the currently available deep learning approaches. The findings are presented in the accompanying table 1:

Table 1 Comparison of results

Methods	PSNR (dB)	SSIM
Dark Channel Prior (DCP)	9.8659	0.5327
IAOD-Net	14.2414	0.6328

By analyzing images, the improved model successfully mitigates haze effects while preserving a greater amount of image intricacies. The subsequent illustration, Figure 2, demonstrates this process:

(a) (b) (c) (d)

Figure 2 A Comparative Analysis of the Performance between DCP and the Newly Proposed IAOD-Net was Conducted Using the Synthetic Indoor Dataset NYU2. (a) Original Image; (b) Represents the Output Image Obtained after Applying the New Approach for Haze Removal; (c) Illustrates the Generated Hazy Picture Derived from (a); (d) depicts the image that has undergone dehazing using DCP[1]

5 DISCUSSION

The improved model of AOD-Net demonstrates strong effectiveness in tasks related to image dehazing, successfully addressing the issues of low brightness and unclear details that are often encountered with conventional techniques. The addition of a multi-scale feature integration module along with an attention generation component greatly enhances the network's adaptability. The findings from the experiments indicate a marked improvement in performance of the updated model compared to the conventional DCP model. This study examined the effectiveness of integrating cross-layer connections within the K-estimation module, allowing for the combination of features across various scales derived from differently sized filters, which aids in achieving a seamless flow from fundamental to advanced features. To enhance the model's adaptability to various data conditions, a modification is made to the standard convolution process by incorporating conditional convolution, which preserves the integration of features across different layers while enabling the network to dynamically modify the convolution kernels based on the specific input. The introduced attention framework amplifies the edge details within the image and incorporates both channel-wise and spatial attention modules into the input tensor, thereby enhancing the model's accuracy and overall performance.

This study introduces an innovative attention mechanism along with a novel conditional convolutional layer to develop a comprehensive end-to-end dehazing neural network. When compared to the classic model employing the same neural network structure, the study achieves a substantially improved objective performance in dehazing on the nyu2 dataset. The research achieves an outstanding Peak Signal-to-Noise Ratio (PSNR) value of 14.Additionally, it achieves an optimal Structural Similarity Index (SSIM) of 0 and a relative enhancement of 44%, equating to 24.6328, indicating an improvement of 18% relative.

6 CONCLUSIONS

This paper introduces a refined model that builds upon the AOD-Net framework, which is an end-to-end CNN designed to effectively generate images without haze. This new approach retains the strengths of the AOD-Net architecture while employing conditional convolution in place of traditional convolution methods. The new model was evaluated against leading techniques using objective measures such as peak signal-to-noise ratio and SSIM on both synthetic and natural haze images. Comprehensive experimental findings illustrate the advantages, reliability, and effectiveness of the new model. The improved model notably achieves better detection capabilities under hazy environments when compared to DCP. Nonetheless, the effectiveness of dehazing methods is closely tied to the accuracy of depth estimation, indicating that integrating depth-related insights or enhanced depth estimation techniques could further elevate the performance of AOD-Net.

Looking ahead, this improved model has the potential to be utilized across various real-world applications, including the enhancement of nighttime imagery in self-driving vehicles. Investigate the potential integration of additional deep learning architectures, like Transformers, to enhance the model's efficiency even further. In subsequent research efforts, the intention is to utilize the complete OTS and ITS datasets to train the proposed system. Next, an in-depth examination of the Aod-Net framework will be undertaken as part of the study. While the present investigation lacked ample opportunity to meticulously adjust the system's parameters, the preliminary findings are promising.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] He K, Sun J, Tang X. Single image haze removal using dark channel prior. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(12): 2341-2353.
[2] Jin Y Z, Chen J X, Tian F, et al. LFD-Net: Lightweight feature-interaction dehazing network for real-time remote sensing tasks. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023, 16: 9139-9153.
[3] Goyal B, Dogra A, Lepcha D C, et al. Recent advances in image dehazing: Formal analysis to automated approaches. Information Fusion, 2024, 104: 102151.
[4] Li Y, et al. Low-light image enhancement via deep learning-based feature attention. IEEE Access, 2020, 8: 179524-179536.
[5] Liu H Z, Hu H M. From global to local: An adaptive environmental illumination estimation for non-uniform scattering//Proceedings of the ACM Multimedia Asia 2023. 2023: 1-7.
[6] Li B, Peng X, Wang Z, et al. AOD-net: All-in-one dehazing network//2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017: 4780-4788.
[7] Mccartney E J. Optics of the atmosphere: Scattering by molecules and particles. New York: Wiley, 1976.
[8] Narasimhan S G, Nayar S K. Contrast restoration of weather degraded images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(6): 713-724.
[9] Nayar S K, Narasimhan S G. Vision in bad weather//Proceedings of the Seventh IEEE International Conference on Computer Vision. Kerkyra: IEEE, 1999: 820-827.
[10] Li B, Ren W, Fu D, et al. Benchmarking single image dehazing and beyond. IEEE Transactions on Image Processing, 2019, 28(1): 492-505.
[11] Gao Z L, Xie F, Yin Y, et al. Image defogging algorithm based on semi-supervised learning neural network//2023 3rd International Conference on Robotics, Automation and Intelligent Control (ICRAIC). 2023: 86-89.
[12] Chen S Y, Cheng J, Huang Z Y. GADO-Net: An improved AOD-Net single image dehazing algorithm//2021 3rd International Academic Exchange Conference on Science and Technology Innovation (IAECST). 2021: 640-646.
[13] ARIF Z H, MAHMOUD M A, ABDULKAREEM K H, et al. Comprehensive review of machine learning (ML) in image defogging: Taxonomy of concepts, scenes, feature extraction, and classification techniques. IET Image Processing, 2021, 16(2): 289-310. DOI:10.1049/ipr2.12365.
[14] Anonymous. Research progress and applications of image defogging algorithms. Progress in Medical Devices, 2023.
[15] Agrawal S C, Jalal A S. A joint cumulative distribution function and gradient fusion based method for dehazing of long shot hazy images. Journal of Visual Communication and Image Representation, 2021, 77: 103087.
[16] Zhang J, Dong Q Q, Song W J. GGADN: Guided generative adversarial dehazing network. Soft Computing, 2021, 27(3): 1731-1741.
[17] Nair D, Sankaran P. Benchmarking single image dehazing methods. SN Computer Science, 2021, 3(1): 1.
[18] Naz B, Rahman N, Nafis M T. Enhancement of dehazing methods using DNN and filtering//Computer Networks and Inventive Communication Technologies. Singapore: Springer, 2021: 531-544.