ARTICLE

Advancements in Remote Sensing Image Dehazing: Introducing URA-Net with Multi-Scale Dense Feature Fusion Clusters and Gated Jump Connection

Hongchi Liu[1], Xing Deng[1,*] and Haijian Shao[1,2]

[1]School of Computer, Jiangsu University of Science and Technology, Zhenjiang, 212003, China

[2]Department of Electrical and Computer Engineering, University of Nevada, Las Vegas, 89119, USA

*Corresponding Author: Xing Deng. Email: xdeng@just.edu.cn

ABSTRACT

The degradation of optical remote sensing images due to atmospheric haze poses a significant obstacle, profoundly impeding their effective utilization across various domains. Dehazing methodologies have emerged as pivotal components of image preprocessing, fostering an improvement in the quality of remote sensing imagery. This enhancement renders remote sensing data more indispensable, thereby enhancing the accuracy of target identification. Conventional defogging techniques based on simplistic atmospheric degradation models have proven inadequate for mitigating non-uniform haze within remotely sensed images. In response to this challenge, a novel UNet Residual Attention Network (URA-Net) is proposed. This paradigmatic approach materializes as an end-to-end convolutional neural network distinguished by its utilization of multi-scale dense feature fusion clusters and gated jump connections. The essence of our methodology lies in local feature fusion within dense residual clusters, enabling the extraction of pertinent features from both preceding and current local data, depending on contextual demands. The intelligently orchestrated gated structures facilitate the propagation of these features to the decoder, resulting in superior outcomes in haze removal. Empirical validation through a plethora of experiments substantiates the efficacy of URA-Net, demonstrating its superior performance compared to existing methods when applied to established datasets for remote sensing image defogging. On the RICE-1 dataset, URA-Net achieves a Peak Signal-to-Noise Ratio (PSNR) of 29.07 dB, surpassing the Dark Channel Prior (DCP) by 11.17 dB, the All-in-One Network for Dehazing (AOD) by 7.82 dB, the Optimal Transmission Map and Adaptive Atmospheric Light For Dehazing (OTM-AAL) by 5.37 dB, the Unsupervised Single Image Dehazing (USID) by 8.0 dB, and the Superpixel-based Remote Sensing Image Dehazing (SRD) by 8.5 dB. Particularly noteworthy, on the SateHaze1k dataset, URA-Net attains preeminence in overall performance, yielding defogged images characterized by consistent visual quality. This underscores the contribution of the research to the advancement of remote sensing technology, providing a robust and efficient solution for alleviating the adverse effects of haze on image quality.

KEYWORDS

Remote sensing image; image dehazing; deep learning; feature fusion

1 Introduction

Haze stands as a prevalent atmospheric phenomenon in nature, predominantly stemming from natural scattering phenomena [1,2]. Its primary constituent is particulate matter suspended within the atmospheric milieu. When the radiative flux from a scene traverses the atmospheric medium and converges upon the imaging apparatus, it undergoes intricate interactions with haze particles, encompassing processes of absorption, reflection, and refraction [3]. Consequently, the resultant images captured by remote sensing equipment are vulnerable to the deleterious consequences of haze, manifesting in perceptible edge blurring and color distortion [4], as illustrated in Fig. 1. This compromise in image fidelity hinders the accurate portrayal of ground surface information, thereby diminishing the efficacy of remote sensing technology. The attenuation in the precision of ground surface information, in turn, curtails the utilization rate and intrinsic value of remote sensing satellites.
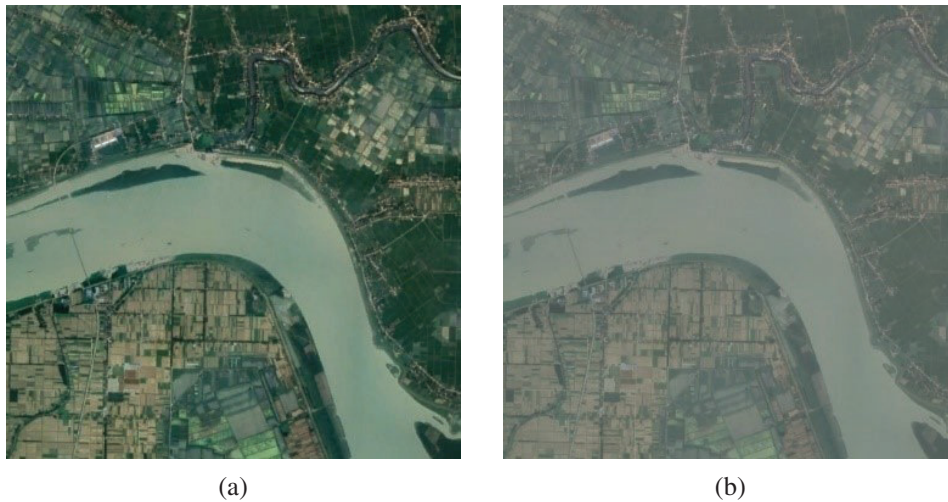

(a)　　　　　　　　　　　　　　　　　(b)

Figure 1: Remotely sensed hazy image and ground truth image. (a) Input hazy image. (b) Ground truth

Currently, models designed for remote sensing image defogging can be broadly classified into three primary categories: image enhancement-based dehazing algorithms, physical model-based defogging algorithms, and data-driven-based algorithms.

1.1 Image Enhancement-Based Dehazing Algorithm

This methodology, facilitated through foggy day image enhancement, endeavors to ameliorate image contrast while minimizing the impact of image noise, thereby effectuating the restoration of a lucid image devoid of fog [5,6]. The inherent advantage lies in the utilization of well-established image processing algorithms that target the enhancement of image contrast [7], thereby accentuating scene characteristics and pertinent information within the image. However, a drawback is discerned in the potential loss of certain information, leading to image distortion. Notwithstanding, this algorithmic approach boasts broad applicability. Representative algorithms within this category encompass the Retinex algorithm [8–12], histogram equalization algorithm [13–17], the partial differential equation algorithm [18,19], the wavelet transform algorithm, and analogous methods [20–23]. While these techniques often yield visually appealing images with relatively efficient computational times, their efficacy is constrained by the inability to comprehensively eradicate haze, given the absence of explicit degradation processes.

The methodology for enhancing images captured under foggy conditions encompasses a suite of techniques, including the Retinex algorithm, histogram equalization algorithm, partial differential equation algorithm, and wavelet transform algorithm. These methodologies are employed to enhance image contrast and reduce noise, thereby restoring clarity to fog-obscured images. Leveraging established image processing algorithms, these techniques aim to emphasize scene characteristics and essential information. Despite their widespread applicability and relatively efficient computational characteristics, a notable limitation resides in the potential loss of information, which can lead to image distortion. Furthermore, these techniques encounter challenges in completely mitigating haze due to the absence of explicit degradation modeling. The selection of a particular technique entails a trade-off between heightened contrast and the risk of distortion, necessitating careful consideration aligned with the specific demands of the application at hand.

1.2 Physical Model-Based Defogging Algorithm

This category of algorithms is predicated on the principles of atmospheric scattering theory, resolving the atmospheric scattering (Eq. (1)) by imbuing it with supplementary conditions to render it nonpathological [24,25]. These algorithms leverage diverse *a priori* information as auxiliary inputs to estimate the parameters λ, η, and β in Eq. (1). Subsequently, by harnessing image degradation, these algorithms enable the direct restoration of the fog-free map [26–28]. It is noteworthy that a common limitation shared by *a priori* algorithms is the circumstance-specific nature of their efficacy; optimal results in *a priori* estimation are attainable only under specific scenario conditions, thereby engendering variable performance across diverse scenarios. The atmospheric scattering equation is articulated as follows:

$$I(x) = J(x)t(x) + A[1 - t(x)], t(x) = e^{-\beta d(x)} \tag{1}$$

where $I(x)$ is the fogged image, $J(x)$ is the reflected light from the observation target, the image after de-fogging, A and $t(x)$ are the atmospheric light coefficient and the atmospheric transmittance, respectively, $d(x)$ is the depth map of the scene and β is the atmospheric light scattering coefficient. It is evident from Eq. (1) that the fog-free image, denoted by $J(x)$ can be recovered from the fogged image $I(x)$ provided that accurate estimations of atmospheric transmittance $((t(x)))$ and atmospheric light (A) are obtained.

1.2.1 Dark Channel Prior (DCP)

He et al. [29] conducted an observation of an extensive *corpus* of outdoor haze-free images, revealing that within the majority of non-sky patches, there exists a noteworthy occurrence wherein at least one color channel manifests pixels with markedly low intensities, proximal to zero, as shown in Fig. 2. For an arbitrary image J its dark channel [30–33] J^{dark} is given by:

$$J^{dark}(x) = \min_{y \in \Omega(x)} \left(\min_{C \in \{R,G,B\}} J^C(Y) \right) \tag{2}$$

where J^C is a color channel of J, and $\Omega(x)$ is a local patch centered at x. The value of J^{dark} should be very low or close to zero if J is an outdoor haze-free image. It is imperative to acknowledge that the diminished intensity observed in the dark channel primarily stems from shadows cast by the scene, the presence of dark objects, and vividly colored surfaces or objects.
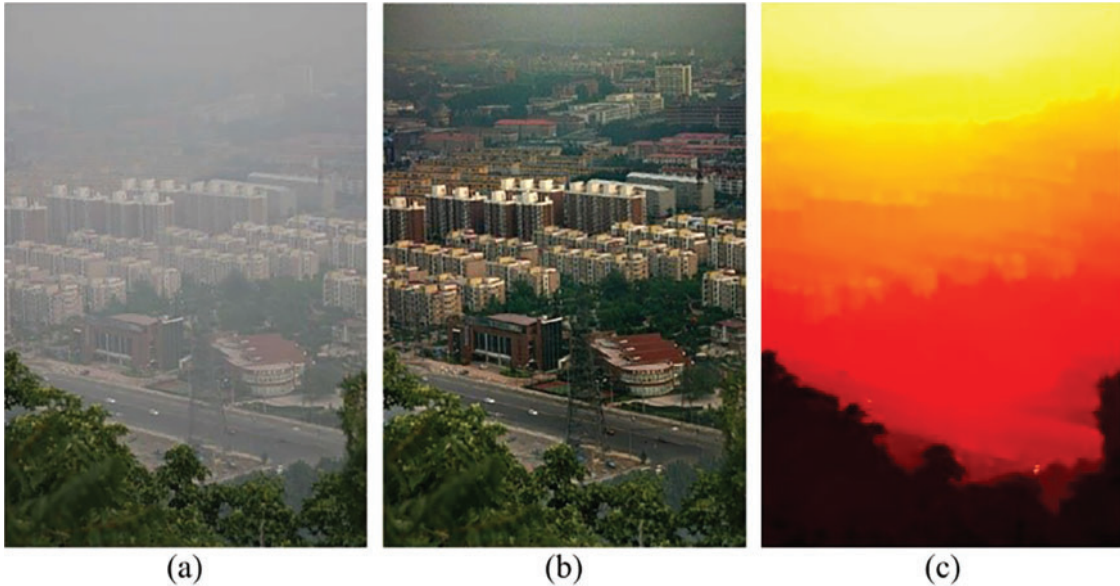
Figure 2: Dark channel removal method [5]: (a) hazy image, (b) dehazed image, and (c) recovered depth-map

The dark channel *a priori* defogging algorithm, grounded in statistical principles, has demonstrated commendable outcomes within the realm of image defogging, exhibiting greater stability in comparison to the aforementioned non-physical model algorithms [34–38]. This algorithm can more accurately estimate the thickness of the fog, resulting in a more natural and clearer defogging effect [39–41]. Nevertheless, a caveat emerges in the form of a discernible darkening effect, and its performance exhibits instability when confronted with foggy images featuring a sky background.

1.2.2 Color Attenuation Prior (CAP)

Through a comprehensive experimental observation of an extensive collection of foggy images, a discernible correlation has been identified between the variance in fog and the brightness as well as saturation of the images. This empirical insight has led to the formulation of a linear model for reconstructing the depth map of the image scene. Subsequently, by leveraging this model, the estimation of atmospheric transmittance $t(x)$ becomes feasible through the utilization of $t(x) = e^{-\beta d(x)}$. The final step involves the restoration of the fog-free image, achieved through the application of Eq. (3).

$$J(x) = \frac{I(x) - A}{\max(t(x), t_0)} + A \tag{3}$$

This category of algorithms exhibits heightened effectiveness in handling mist images characterized by a distinct and unobstructed background, coupled with significant color saturation. However, a notable susceptibility arises in the form of fog residue phenomena within regions of varying depth of field, thereby diminishing its efficacy. Furthermore, its effectiveness is notably diminished when applied to fog-containing images featuring clustered elements [42–45].

1.3 Data-Driven Based Dehazing Algorithm

In tandem with the relentless evolution of deep learning theory, the Convolutional Neural Network (CNN) [46–49] has garnered widespread adoption and commendable success in diverse domains such as face recognition and image segmentation. CNN networks can efficiently learn complex feature representations from large amounts of image data to achieve excellent performance in visual tasks [50–53]. Image dehazing, a salient concern within the realm of image processing, has garnered considerable scholarly interest. Notably, data-driven dehazing techniques, rooted in deep learning paradigms, have demonstrated remarkable advancements when juxtaposed against conventional haze removal methods.

Deep learning defogging algorithms can be stratified into two overarching categories: estimated parameter methods and direct repair methods. The parameter estimation method employs deep convolutional neural networks to directly estimate key parameters such as $t(x)$ and A, etc., subsequently facilitating image dehazing through the application of the atmospheric scattering model. Leveraging the learning capability inherent in deep neural networks, the direct inpainting method achieves an end-to-end process, directly deriving the haze-free image from the foggy image within the framework of data-driven conditions.

1.3.1 Parameter Optimization-Based Methods

This approach involves parameter estimation through deep learning techniques, with the parameters $t(x)$ and A, among others, estimated through both non-deep learning physical model algorithms and deep learning estimation parameter methods. Notably, parameters estimated via deep learning exhibit a higher degree of accuracy compared to their non-deep learning physical model counterparts, resulting in a more pronounced improvement in de-fogging effectiveness [54–58]. Parameter estimation through deep learning methods allows for greater flexibility in adapting to a variety of complex scenarios and data distributions, resulting in higher levels of performance improvement [59–63].

However, it is essential to acknowledge that the dehazing method predicated on neural network parameter estimation shares similar drawbacks with physical model dehazing algorithms. Specifically, inaccuracies in parameter estimation significantly compromise the efficacy of dehazing. Additionally, since the neural network's parameters are trained on synthetic datasets, the challenge arises in truly capturing the relevant characteristics of real environmental fog. Consequently, the application of such algorithms to dehaze real scenes may lead to a degradation in defogging quality.

1.3.2 Advancements in Dehazing for Enhanced Image Restoration

To circumvent the constraints associated with estimated parameters, researchers have harnessed the formidable learning capabilities of neural networks, ushering in a new wave of direct repair dehazing algorithms [64–68]. This new approach shows high robustness and generalization ability when dealing with various types and degrees of haze images [69–72]. This category can be further classified into feature fusion dehazing and adversarial generation dehazing algorithms.

Gated Context Aggregation Network (GCA-Net) [73], leveraging smoothed dilation convolution and threshold fusion, excels in producing commendable results by enabling the network to learn residuals between $J(x)$ and $I(x)$. The algorithm constructs a loss function and undergoes training to optimize network performance. Gated Fusion Network (GFN) [74], achieving noteworthy outcomes through white balancing, contrast enhancement, and λ correction for foggy maps, employs three auxiliary maps as inputs to the network, learning belief maps of White Balance (WB), Contrast Enhancing (CE),

and Gamma Correction (GC). The algorithm, despite exhibiting superior performance compared to parameter estimation algorithms, is characterized by a heightened complexity.

Grid-Dehaze Net [75] achieves enhanced results by incorporating a multi-scale attention mechanism through the utilization of a grid network in its backbone. Feature Fusion Attention Network (FFA-Net) [76], introducing both channel attention and spatial attention mechanisms, achieves superior outcomes through the integration of a learning structure containing the attention mechanism and global residual connections. Notably, this algorithm achieves remarkable results with L1 loss alone. The efficacy of these algorithms in de-fogging is contingent upon the specific feature fusion methods employed, necessitating meticulous design for optimal results.

Increasing the visibility of nighttime hazy images is challenging because of uneven illumination from active artificial light sources and haze absorbing/scattering. To address this issue, Zhang et al. [77] proposed a novel method to simulate nighttime hazy images from daytime clear images, and finally render the haze effects. Degradation-Adaptive Neural Networks can adapt to different types and levels of image degradation. Chen et al. [78] designed a high-quality single-image defogging algorithm based on a degenerate adaptive neural network and achieved good results. The scheme, which is called Towards Perceptual Image Dehazing (TPID) [79] leverages three generative adversarial networks (GJ, GA, and Gt indicate the generators for the scene radiance, the medium transmission, and the global atmosphere light, respectively) for image de-fogging, employing an atmospheric scattering model. It addresses the inherent dependency on pair-wise datasets in deep learning de-fogging by constructing unpaired fogged and unfogged datasets through a combination of adversarial loss and L1 loss. ICycleGAN [80] integrates the concept of Cycle Generative Adversarial Network (GAN), incorporates perceptual loss L Perceptual, and undergoes network training. While these algorithms exhibit efficacy in real scene de-fogging, they are susceptible to dataset dependence and prone to color shift distortion issues. In this context, Fig. 3 illustrates a comparison of URA-net with other image dehazing methods on the RICE-I set. Inspired by the aforementioned developments, this paper proposes a U-Net-based end-to-end attention residual network, denoted as URA-Net, featuring an updated encoder-decoder structure. The key contributions of this study are summarized as follows:

1. Residual structures are introduced to the backbone to alleviate the gradient vanishing problem, enhancing the generalization performance of the backbone. A simple attention module is incorporated into the residual unit to bolster the trunk's capacity for extracting fine-grained features, thereby improving defogging performance.

2. Jump connections are further enhanced by integrating a gating mechanism, allowing attention coefficients to be more precisely targeted to localized regions, resulting in improved overall performance.

The subsequent sections of the paper are organized as follows: Section 2 presents an overview of related work; Section 3 elucidates the proposed image-defogging framework; Sections 4–6 provide a discussion and comparative analysis of experimental results; and Section 7 concludes the paper.
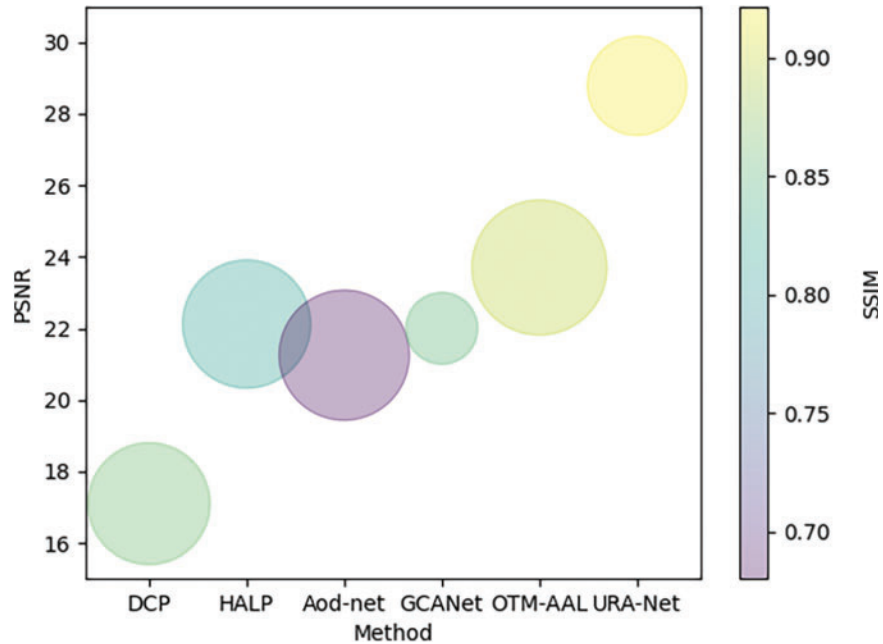
Figure 3: Comparison of URA-net with other image dehazing methods on the RICE-I set

2 Related Work

2.1 Remote Sensing Image Defogging-Related Approaches

The quest to mitigate atmospheric haze in remote sensing images and enhance their quality has given rise to two principal categories of methods: traditional *a priori* knowledge-based approaches and deep learning-based methods.

Traditional *a priori* methods typically commence by estimating transmittance maps for atmospheric light and hazy images, followed by the application of classical atmospheric scattering models to reconstruct fog-free images. However, these methods heavily hinge on assumptions about the scene and haze properties, rendering their defogging performance susceptible to inaccuracies in these assumptions [81–83]. Additionally, these approaches are formulated based on diverse assumptions and models, constraining their universal applicability across various environmental conditions and atmospheric compositions.

In contrast, deep learning-based methods for de-fogging remote sensing images leverage advanced algorithms to learn the mapping between fogged and fog-free images. Utilizing extensive datasets containing paired foggy and clear images, these methods have demonstrated efficacy in alleviating haze in diverse remote sensing scenarios.

2.2 Deep Learning-Based Approach for De-Fogging

Deep learning-based image-defogging methods can be broadly categorized into two main groups. These categories are determined by the architectural design of the defogging networks, with commonly employed networks either operating independently of atmospheric scattering models or incorporating

them. The former category includes encoder-decoder networks, GAN-based networks, attention-based networks, and transformer-based networks. Notably, these state-of-the-art networks demonstrate superior generalization capabilities and robustness compared to their early counterparts, making them the mainstream methods in the field. Despite their advancements, the complex architectures of these networks pose challenges. While they enhance performance, they also significantly increase network complexity, introducing challenges in terms of training, inference, and deployment [84,85]. Balancing performance improvement and complexity is an ongoing area of research in de-fogging networks, addressing the challenge this paper seeks to resolve. These methods contribute to reducing the defogging network's reliance on specific datasets, endowing them with more robust generalization abilities.

2.3 U-Net-Related Approaches

Conventional neural networks (CNNs) demonstrate proficiency in learning multi-level image features but encounter difficulties in achieving pixel-level image classification. The introduction of Fully Convolutional Networks (FCN) represented a significant leap forward in image segmentation compared to traditional CNNs. FCN achieves end-to-end image segmentation by up-sampling the feature mapping from its final convolutional layer [86,87]. However, the relatively simplistic structure of FCN imposes limitations on its overall performance. Addressing this constraint, U-Net enhances image feature fusion by introducing an encoder-decoder structure based on FCN. The encoder module of U-Net utilizes convolution and down-sampling to extract shallow features from the image [88–91]. Simultaneously, the decoder module employs deconvolution and up-sampling to capture deep image features. U-Net further refines segmentation by incorporating skip connections between shallow and deep features, facilitating the extraction of finer image details.

Despite these advancements, an increase in the number of network layers often results in a performance decline due to the challenge of gradient vanishing. To tackle this issue, He et al. proposed the widely acclaimed Residual Network (ResNet). Res-Net adopts a novel approach by focusing on learning the residual representation between inputs and outputs rather than solely emphasizing feature mapping [92,93]. This architectural innovation has proven effective in mitigating the impact of gradient vanishing, leading to improved network performance, especially in tasks involving a large number of layers.

3 The Proposed URA-Net for Remote Sensing Image Dehazing

Fig. 4 delineates the comprehensive architecture of the proposed URA-Net, showcasing its evolution through the synthesis of U-Net and Res-Net principles. Notably, URA-Net preserves the foundational encoder-decoder structure of U-Net while introducing an additional seven building blocks at the junction, referred to as the bottleneck structure, strategically positioned between the encoder and decoder modules. Moreover, URA-Net incorporates the fundamental concept of Res-Net, emphasizing the learning of residual images rather than direct image defogging. This is realized by introducing the residual image and integrating it with the original blurred image, culminating in the derivation of the final image-defogging result. The amalgamation of U-Net's encoder-decoder architecture with Res-Net's residual learning approach empowers URA-Net to effectively address atmospheric haze in remote sensing images, as visually depicted in the proposed architectural representation.
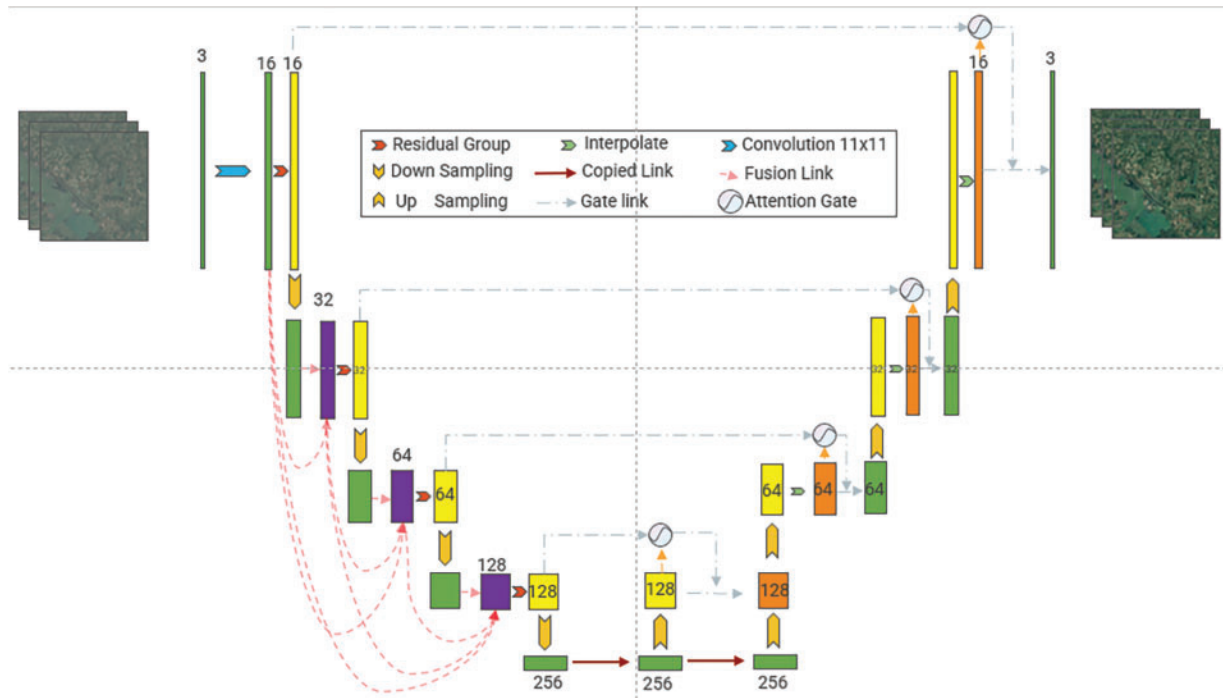
Figure 4: The architecture of URA-Net

The architecture of URA-Net, as illustrated in Fig. 4, encompasses three principal components: an encoder module, a jump connection (as referenced in [94,95]), and a decoder module. Specifically, the encoder module is engineered to extract and amalgamate the feature mappings of an image via a dense feature fusion module. Within URA-Net, the primary objective of the decoder module is to acquire the depth feature mapping of the image through inverse convolution operations, a process involving the reconstruction of detailed features from the encoded representation. Moreover, to augment feature fusion and facilitate information propagation, a skip connection is established between the encoder module and the decoder module. This connection enables the integration of shallow feature mappings from the encoder with deep feature mappings from the decoder. Fundamentally, URA-Net adopts an encoder-decoder architecture to learn the residual image, as discussed in [96,97]. The ultimate defogging outcome of the image is derived by adding the original blurred image to the computed residual image. This methodology not only captures intricate details through the decoder's inverse convolution but also ensures the retention of crucial information from both shallow and deep feature levels, thereby enhancing the overall efficacy of the image-defogging process. The central operational mechanism of URA-Net is elucidated in Fig. 5.
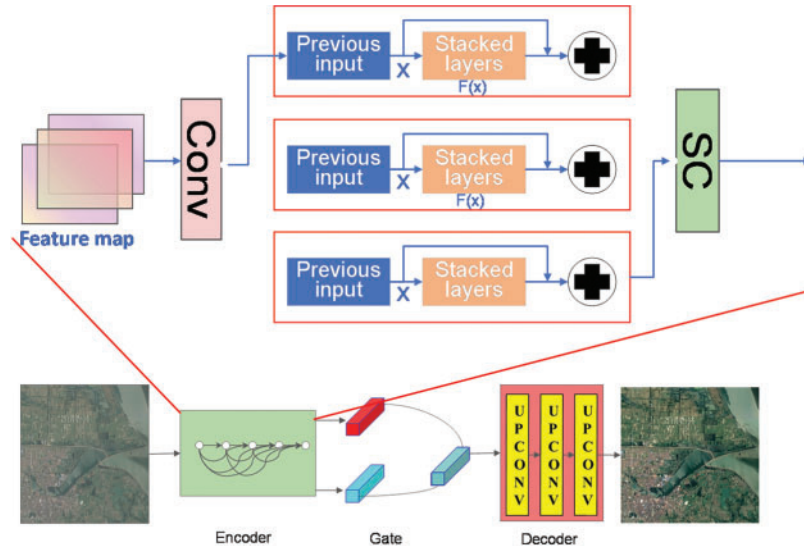
Figue 5: Schematic flow diagram of URA-Net

3.1 Residual Group

This paper predominantly employs a neural network model based on PyTorch, leveraging key building blocks to enhance the network's performance. The 'make dense' class defines a dense connection block, fostering feature accumulation by linking input and output in the channel dimension. This design facilitates information transmission and gradient flow, mitigating the issue of vanishing gradients. The introduction of residual blocks, instantiated through the 'ResidualBlock' class, further bolsters gradient propagation. These blocks, connected via skips, enable direct gradient flow to shallow layers, alleviating vanishing gradient challenges in deep networks–a crucial aspect for effectively capturing complex image features.

The overall network structure follows a U-shaped encoding-decoding architecture. The encoding section captures image features, while the decoding portion progressively restores spatial resolution through up-sampling convolution. The *'UpsampleConvLayer'* class handles the upsampling operation, with interpolation implemented using PyTorch's *'F.interpolate'* function. In the decoding phase, the *'AttentionBlock'* class introduces an attention mechanism, enabling the model to focus on crucial feature areas during image reconstruction [98–101]. This inclusion enhances the network's performance and generalization capabilities. Batch normalization and the Rectified Liner Units (RELU) activation function are integral components of the code, contributing to improved stability and faster convergence of the network. The entire model is instantiated on Compute Unified Device Architecture (CUDA) devices and leverages GPU acceleration, underscoring its commitment to efficient processing. In summary, this paper's neural network model adeptly combines dense connections, residual blocks, multi-scale design, and attention mechanisms, creating an image-processing network with robust generalization capabilities. The U-Net architecture exhibits inherent limitations, particularly in spatial information loss during encoder down-sampling and insufficient connectivity between features from non-adjacent levels. To address these drawbacks, a straightforward approach involves resampling all features to a common scale and subsequently fusing them through bottleneck layers, incorporating connectivity and convolutional layers, similar to DenseNet [102]. However, the conventional splicing method for feature fusion proves less effective, primarily due to variations in scales and dimensions

across features from different layers. In response to these challenges, we propose a novel Dense Feature Fusion with Attention (DFFA) module to efficiently compensate for missing information and leverage features from non-adjacent layers [103,104]. The DFFA module is designed to enhance features at the current level through an error feedback mechanism, specifically implemented in the encoder.

As depicted in the architecture, a DFFA module is strategically introduced before the residual group in the encoder for effective feature extraction and fusion. This module not only addresses the limitations of the U-Net architecture but also leverages an error feedback mechanism to enhance the features at each level, ensuring improved information compensation and connectivity across non-adjacent layers [105,106]. The utilization of the DFFA module represents a significant enhancement to the overall network, contributing to the efficiency and efficacy of feature extraction and fusion processes.

3.2 Dense Feature Fusion

Following the extraction of local dense features utilizing a series of Dense Blocks, the introduction of DFFA aims to encapsulate hierarchical features from a global standpoint. The DFFA is structured around two principal constituents: global feature fusion (GFF) and global residual learning, as referenced in [107,108]. Within this framework, global feature fusion entails the derivation of the global feature (FGF) by amalgamating features originating from all Dense Blocks.

$$F_{GF} = H_{GFF} ([F_1 \cdots , F_D]) \tag{4}$$

where $[F_1, \cdots , F_D]$ is the stitching of the feature maps generated by the Dense Block, and H_{GFF} represents the composite function of the 1×1 and 3×3 convolutions.

3.3 Global Residual Learning

The feature mappings are then obtained through global residual learning, followed by Eq. (5), and scaling is then performed

$$F_{DF} = F_{LF} + F_{GF} \tag{5}$$

where F_{LF} denotes shallow feature mapping. Our proposed residual dense block makes full use of all other layers before global feature fusion. The Residual Group produces multi-level local dense features, which are further adaptively fused to form F_{GF}. After global residual learning, dense features are obtained as F_{DF}.

3.4 Attention Gate

Algorithm 1: Pseudo code of attention gate

Require: g: Feature map, x: Feature map
Ensure: $\hat{x}$: Fusion feature map
  1: $g' \leftarrow W_g (g)$
  2: $x' \leftarrow W_x (x)$
  3: $\sigma_1 \leftarrow Add (g' + x')$
  4: $\sigma_2 \leftarrow Relu (\sigma_1)$
  5: $\zeta_1 \leftarrow \psi (\sigma_2)$
  6: $\zeta_2 \leftarrow Sigmond (\zeta_1)$
  7: $\alpha \leftarrow Resample (\zeta_2)$
  8: $\hat{x} (Fusion feature map) \leftarrow \alpha \times x$

Attention gates [109,110] are akin to a human visual attention mechanism that automatically focuses on the target region. They learn to suppress irrelevant feature responses in the feature map while highlighting salient feature information crucial for a particular task. From the pseudo-code flow in Algorithm 1, it is evident that the inputs of the Attention gate are two feature maps (g, x) These correspond to the feature map of the decoder part and the feature map of the encoder part of the layer above it. The specific steps are given as follows: firstly, g and x are operated in parallel, and g' is calculated by means W_g of g and x' is then derived by means W_x of. Subsequently, the operation $g + x$ is thus performed to obtain σ_1. Then σ_2 is calculated through the Rulu operation on σ_1, ζ_1 is calculated by means ψ of σ_2, ζ_2 is calculated through means of ζ_1, and α is then derived by means of. Finally, multiply α by x to get the fusion feature. A typical architecture of an attention-gating unit is shown in Fig. 6.
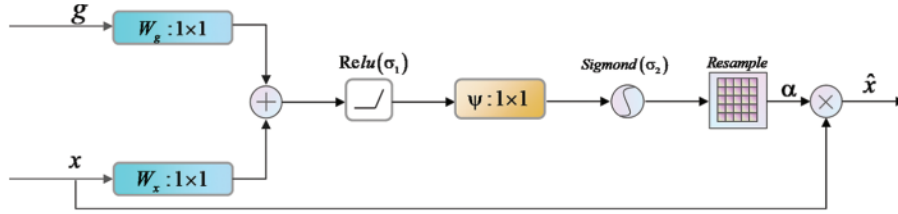

Figure 6: Attention gate control unit

3.5 Loss Function

The loss function of a network plays a crucial role in determining the training efficiency and effectiveness of the network. The mean square error (MSE) loss function is commonly used for image-defogging tasks. However, defogged images obtained using the MSE loss often exhibit blurriness, leading to a mismatch in visual quality. The L1 loss function, being less sensitive to outliers than the MSE loss, has been found to achieve better performance [111].

To address the blurriness issue and enhance the visual quality of defogged images, we introduce the Structural Similarity Index Metric (SSIM) loss. The total loss function is defined as follows:

$$loss = \lambda_1 \times Smooth_{L1}\_LOSS + \lambda_2 \times SSIM\_loss \tag{6}$$

where $\lambda_1 = \lambda_2 = 0.5$, the first L1 smoothing error is based on the MSE, which addresses the zero non-smoothing problem. It is employed to measure the loss between the residual images obtained by URA-Net and the actual residual images. This smoothing mean square error is defined as:

$$Smooth_{L1}(x) = \begin{cases} 0.5x^2 & if\ |x| < 1 \\ |x| - 0.5 & otherwise \end{cases} \tag{7}$$

SSIM is proposed to assess the structural similarity between two images by extracting structural information, which can be defined as follows [112,113]:

$$SSIM(I_{de}, I_{gt}) = \frac{2\mu_{Ide}\mu_{Igt} + C_1}{\mu^2_{Ide} + \mu^2_{Igt} + C_1} \frac{2\sigma_{IdeIgt} + C_2}{\sigma^2_{Ide} + \sigma^2_{Igt} + C_2} \tag{8}$$

where μ_{Ide} and μ_{Igt} are the mean of the defogged and clear images, respectively. σ^2_{Ide} and σ^2_{Igt} are the variance of the defogged and clear images, respectively. σ_{IdeIgt} is the covariance of the defogged and clear images; and C1 and C2 are constants. Correspondingly, the SSIM loss can be expressed as:

$$L_S = 1 - SSIM\left(I_{de}, I_{gt}\right) \tag{9}$$

4 Experiments

We conducted evaluations on the following datasets using state-of-the-art methods:

(i) RICE-I dataset [114]: Contains 500 pairs of images, each with and without clouds, of size 512×512. Randomly selected 400 pairs of images for training and 100 pairs for testing.

(ii) SateHaze1k dataset [115]: Comprises pairs of aerial images, including clean and degraded images with varying levels of haze density (thin, moderate, and thick). Utilizes data augmentation techniques such as random flipping. Consists of 1920 image pairs (degraded and ground-truth) for training, containing light, medium, and dense haze. Includes 45 image pairs dedicated for testing in each level of haze density.

(iii) Competitors & Metrics: We compare our method with several state-of-the-art dehazing algorithms, including Dark Channel Prior (DCP) [29], Dehazing Scheme Using Heterogeneous Atmospheric Light Prior(HALP) [116], All-in-One Network for Dehazing (AOD-Net) [117], FFA-Net [118], GCA-Net [119], the Optimal Transmission Map and Adaptive Atmospheric Light For Dehazin (OTM-AAL) [120], Unsupervised Single Image Dehazing Network (USID-Net) [121] and Uperpixel-Based Remote Sensing Image Dehazing (SRD) [122]. Evaluation metrics include Peak Signal-to-Noise Ratio (PSNR) and SSIM [57] to assess the quality of restored images.

(iv) Implementation details:

(a) In the training phase, we apply the discriminator proposed in [103] with a patch size of 30×30

(b) The Adam optimizer with a learning rate lr $= 10^{-4}$ is used to optimize the network.

(c) A batch size of 16 is employed for training.

4.1 Evaluation Metrics

For hazy images with ground truth, we use the reference image as the ground truth, and the defogging effect can be quantitatively assessed by calculating the statistical values of the image before and after defogging. The metrics mean square error (PSNR) and structural similarity index (SSIM) are employed. PSNR, which stands for peak signal-to-noise ratio, indicates the quality of the images. The higher the PSNR value, the lower the distortion between the two images and the higher the quality. SSIM, or structural similarity index, measures how similar the two images are; the closer SSIM is to 1, the more similar the images are in terms of structure.

For hazy images without ground truth, we introduce BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator) [123], NIQE (Natural Image Quality Evaluator) [124], and PIQE (Perception-based Image Quality Evaluator) [125] as performance evaluation metrics. BRISQUE is an image quality assessment metric based on the analysis of local image features. The NIQE algorithm employs features that are more sensitive to regions with higher contrast in images, without relying on any subjective assessment scores. PIQE is a perception-based image quality assessment algorithm that

frames the problem of image quality evaluation as simulating human visual perception. Additionally, all three evaluation metrics exhibit the characteristic that lower values indicate better image quality.

4.2 Quantitative Comparison

Table 1 presents the results of a comprehensive quantitative comparison of the SateHaze1k and RICE datasets. The proposed URA-Net model demonstrates a noteworthy improvement over other mainstream state-of-the-art methods listed in the table. The performance gains are particularly evident across various metrics. Notably, URA-Net outperforms the second-place method by a margin of 0.1 in terms of SSIM on the dense haze dataset, showcasing a substantial advantage over CNN-based methods. On the RICE dataset, URA-Net exhibits a remarkable improvement, surpassing GCA-Net by more than 2.0 and 8.5 dB over SRD. This compelling performance differential underscores the efficiency and accuracy of our proposed method in handling diverse types of haze distributions. The results affirm URA-Net's robustness and effectiveness in comparison to contemporary approaches.

Table 1: Average SSIM/PSNR on dataset

Methods	Goal	DCP	HALP	Aod-net	GCA-Net	OTM-AAL	USID-Net	SRD	URA-Net
RICE-1	PSNR ↑	17.10	22.12	21.25	22.00	23.70	21.07	20.55	29.07
	SSIM ↑	0.8613	0.8132	0.6875	0.8501	0.8965	0.8371	0.9153	0.9215
RS-Thin	PSNR ↑	20.51	18.47	18.74	19.10	17.60	19.98	21.33	24.01
	SSIM ↑	0.8733	0.8513	0.8460	0.8453	0.8255	0.8453	0.8960	0.9094
RS-Moderate	PSNR ↑	21.36	18.46	19.64	19.56	16.92	20.06	20.77	26.57
	SSIM ↑	0.8942	0.8527	0.8253	0.8420	0.8148	0.8406	0.9300	0.9425
RS-Thick	PSNR ↑	16.39	14.04	15.47	16.62	10.07	16.17	17.26	23.92
	SSIM ↑	0.7535	0.7135	0.6992	0.6981	0.5229	0.7306	0.8140	0.8624

In Table 1, quantitative evaluations on the benchmark dehazing datasets. Red texts and blue texts indicate the best and the second-best performance, respectively. ↑ and ↓ mean the better methods should achieve higher/lower scores of this metric.

Table 2 presents the results of a comprehensive quantitative comparison on the Real-World dataset. Red texts and blue texts indicate the best and the second-best performance respectively. For quantitative comparison, three well-known no-reference (BRISQUE, NIQE and PIQE) image quality assessment indicators are adopted for evaluations. All these metrics are evaluated on 350 real-world hazy images selected from the Goole and public datasets. The BRISQUE, NIQE, and PIQE metrics are used to assess the overall quality of the restored images, with lower values indicating better results. As observed, URA-Net achieves two first and third places in these three indicators, respectively, showing that the haze-free images restored by our method are of high quality.

Table 2: Average BRISQUE/NIQE/PIQE on real-world dataset

	Goal	DCP	HALP	Aod-net	GCA-Net	OTM-AAL	USID-Net	SRD	URA-Net
BRISQUE	↓	31.54	32.03	36.84	34.69	31.69	33.14	34.55	31.35
NIQE	↓	9.975	7.658	8.376	8.003	9.353	7.342	8.614	7.983
PIQE	↓	30.92	28.03	30.11	29.26	31.59	31.15	31.11	27.77

4.3 Qualitative Evaluations

Figs. 7–10 depict the denoising outcomes obtained from RICE-1 and SateHaze1k images, respectively, following resizing to a 512×512 dimension to enhance detail visibility. While DCP and Aod-Net effectively eliminate the translucent cloud layer, they concurrently induce a perceptible darkening across the resultant images. Although GCA-Net exhibits improved fidelity in restoring original colors, its efficacy in deblurring remains wanting, particularly discernible in scenarios where denoising performance falls short. In contrast, OTM-AAL showcases superior deblurring capabilities, albeit at the expense of noticeable image distortion, overexposure in bright regions, and inadequate de-fogging performance. HALP demonstrates enhanced image clarity; however, it comes at the cost of conspicuous exposure effects.

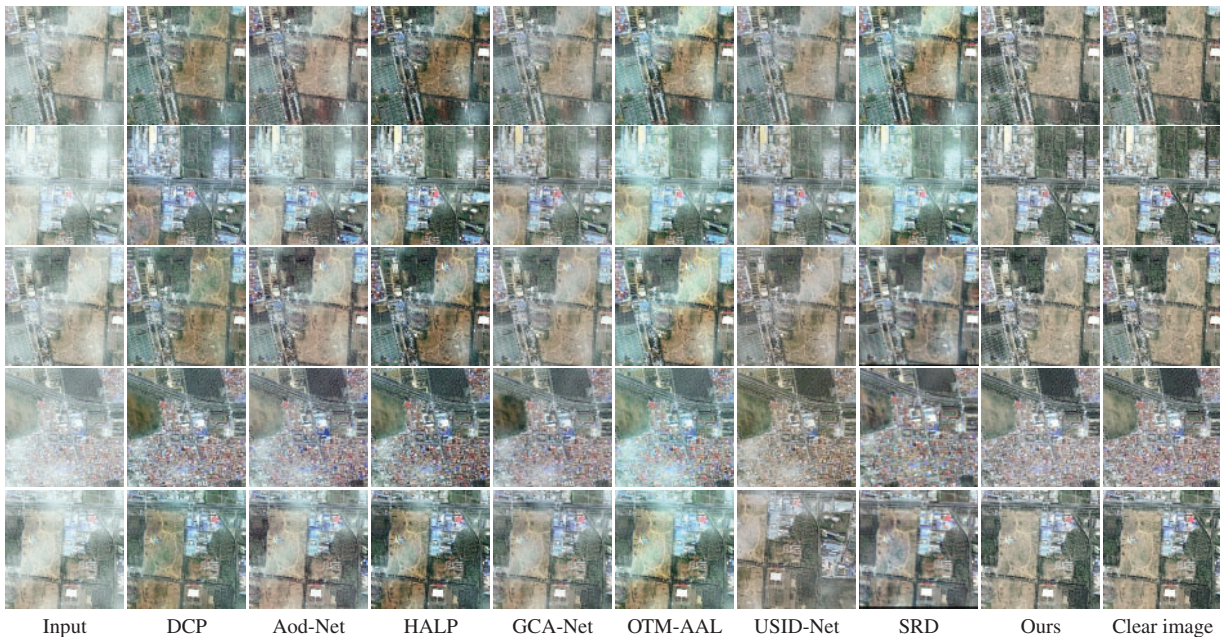Input DCP Aod-Net HALP GCA-Net OTM-AAL USID-Net SRD Ours Clear image

Figure 7: Qualitative comparisons of different methods on Haze1k-Thin

The proposed URA-Net method addresses these inherent shortcomings by assimilating DCP's superior deblurring capabilities while ameliorating the issue of darkened imagery. URA-Net further enhances cloud shadow and dark area details, resulting in a more lucid depiction. In comparison, our algorithm surpasses existing methodologies by furnishing images endowed with sharper structures and details closely approximating ground truth. The perceptual evaluation of results underscores the superior performance and enriched visualization achieved by our proposed URA-Net method.
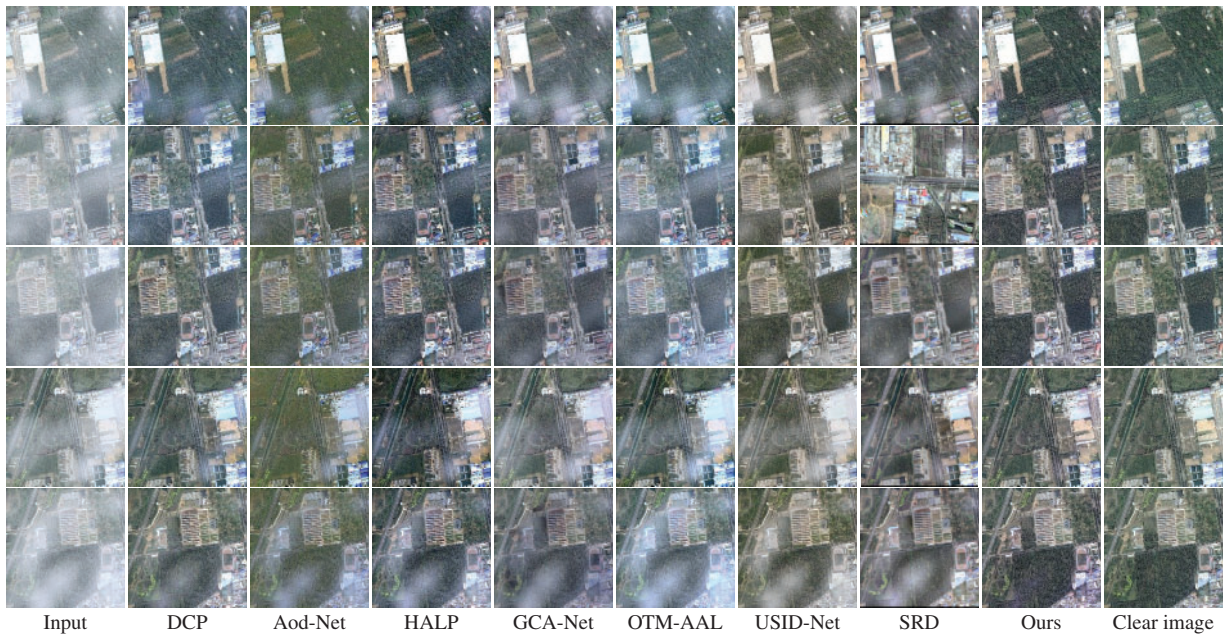
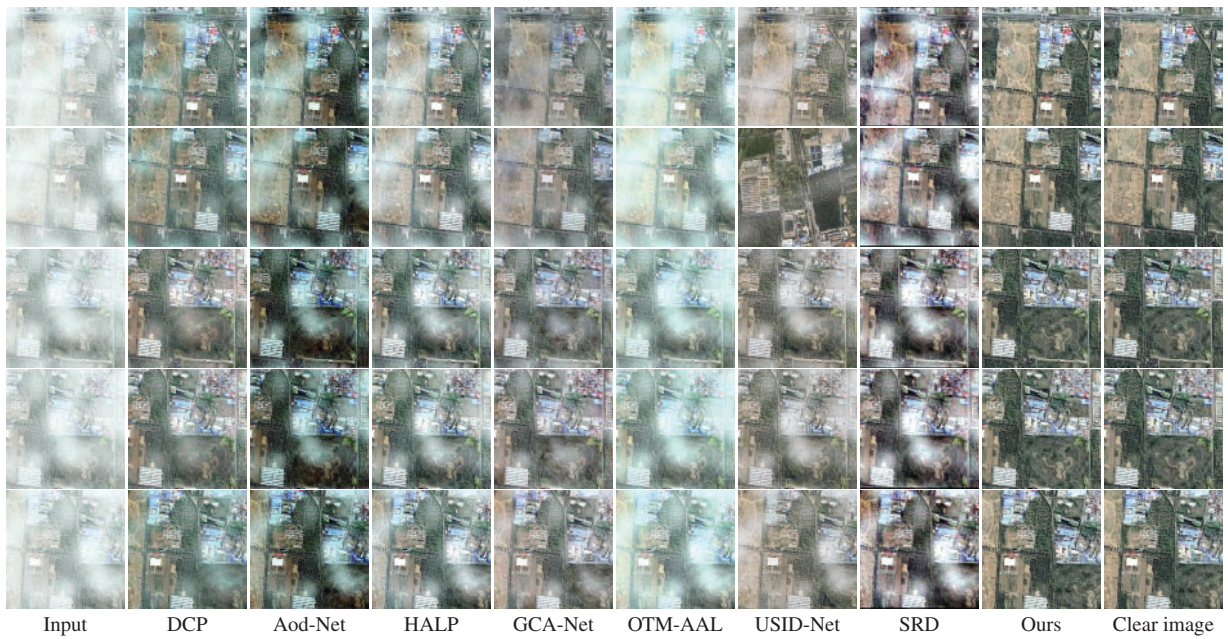**Figure 8:** Qualitative comparisons of different methods on Haze1k-Moderate

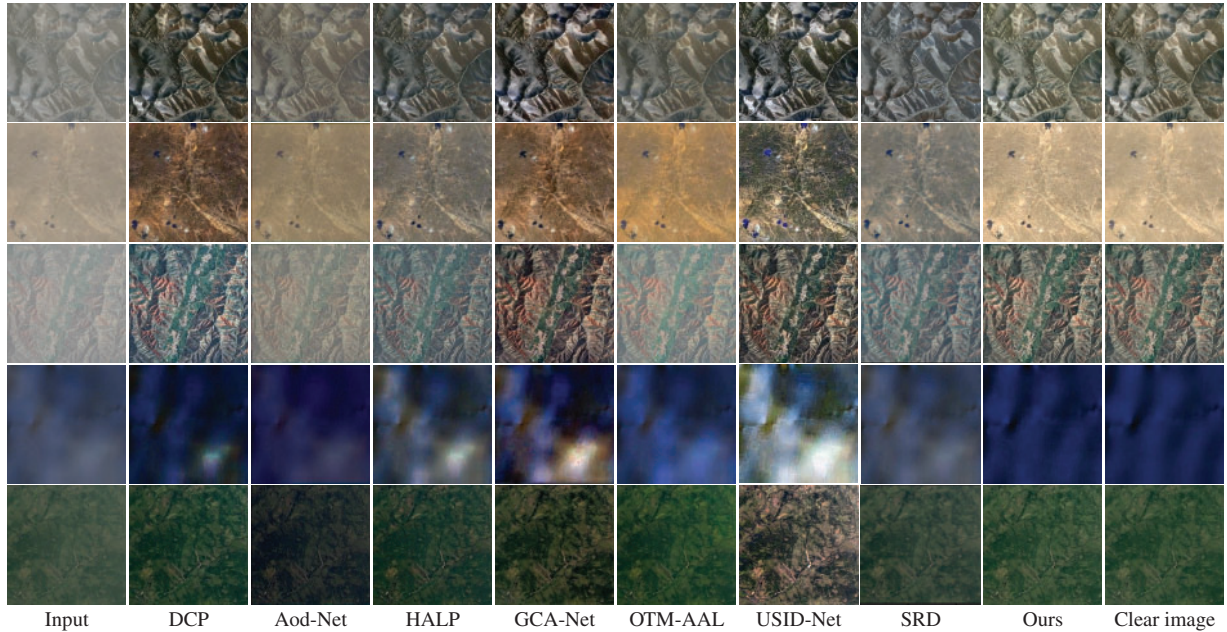**Figure 9:** Qualitative comparisons of different methods on Haze1k-Thick

Figure 10: Qualitative comparisons of different methods on RICE-I dataset

4.4 Results on Real-World Hazy Images

We qualitatively evaluate URA-Net with other dehazing approaches on several real-world hazy images from the Goole-net. Fig. 11 exhibits 4 real-world hazy samples and the corresponding restored results by different dehazing algorithms. As can be seen from the figure, DCP suffers from color distortion (e.g., sky component), which leads to inaccurate color reconstruction. For HALP, AOD, and GCA, the details of objects are relatively blurred in real foggy conditions, probably because these methods do not accurately estimate the fog component when dealing with complex real scenes. OTM-AAL and SRD produce bright effects but have poor and limited de-fogging capabilities. Our method generates better perceptual results in terms of brightness, colorfulness, and haze residue compared to other methods.
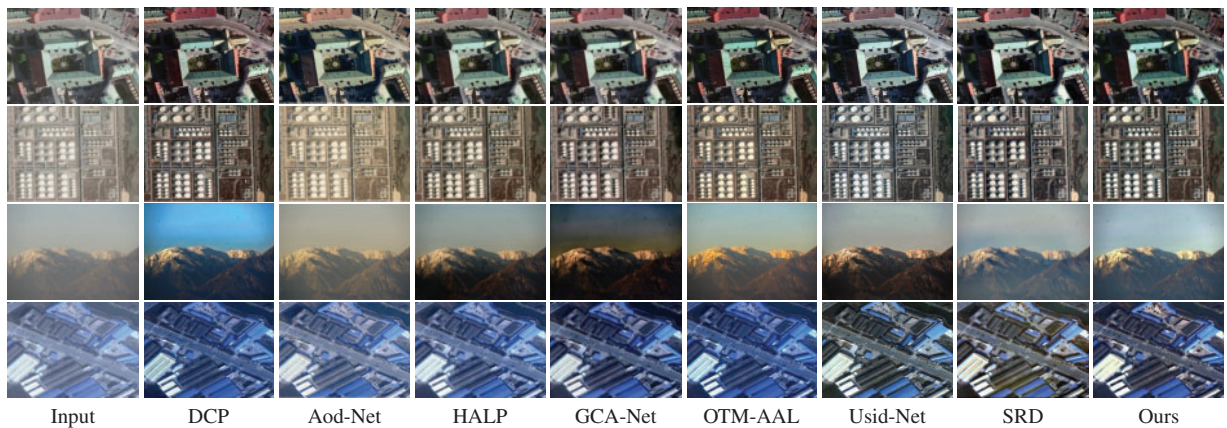

Figure 11: Qualitative comparisons of different methods on real-world dataset

5 Ablation Study

The efficacy of the proposed modules undergoes rigorous evaluation through ablation studies conducted on the StateHaze1K and Real-World datasets. These studies are designed to systematically dissect the URA-Net architecture by isolating and scrutinizing individual components, aiming to discern their specific contributions and impact on overall performance. The overarching objective is to juxtapose these components against counterparts with analogous designs, thereby elucidating their roles and effectiveness within the URA-Net framework.

This methodical approach facilitates a comprehensive comprehension of how each module influences the model's overall efficacy. Through the conduct of ablation studies, the endeavor is to quantitatively gauge and appraise the significance of each proposed component, fostering a nuanced comparison with analogous designs. Such endeavors contribute to a refined understanding of the pivotal elements inherent in the URA-Net architecture and their respective contributions to its performance. We designed five sets of experiments: (a) Without Stacked Cross(SC)-Attention, remove SC-Attention modules from the residual group (b) Without Gate, use common jump connections instead of gated structures (c) Without Dense Feature Fusion, use down-sampling instead of dense connection fusion (d) Without Loss, use MSE-loss instead of our proposed loss and (e) URA-Net, complete network proposed.

Table 3 demonstrates notable enhancements in PSNR metrics, showcasing an elevation of 0.96 with the incorporation of the SC attention mechanism, 0.99 with the introduction of the self-attention gating mechanism, and a substantial increase of 1.42 attributable to the adoption of the devised combined loss function. The amalgamation of these modules is intricately designed to afford discerning attention to both multi-channel and multi-pixel aspects, concurrently facilitating the retention and dissemination of information across diverse network layers.

Table 3: Qualitative comparisons of different variants of URA-Net on the Remote Sensing-Thick

Model	PSNR 2	SSIM
Without SC-Attention	22.96	0.8514
Without gater	22.93	0.8419
Without loss (Only MSE loss)	22.50	0.8310
Without dense feature fusion	21.71	0.8285
URA-Net	23.92	0.8624

Table 4 demonstrates the significant reduction of BRISQUE, NIQE, and PIQE metrics, using the proposed gating mechanism, BRISQUE decreased by 1.32, NIQE decreased by 0.35 and PIQE decreased by 0.80. Due to the use of the designed combined loss function, BRISQUE decreased by 1.86, NIQE decreased by 0.49, and PIQE decreased by 1.66. The experimental results show that the use of each component is beneficial to improve the image quality of real foggy maps after defogging, and verifies the effectiveness of the proposed defogging network in real defogging.

Table 4: Qualitative comparisons of different variants of URA-Net on the Real-World dataset

Model	BRISQUE	NIQE	PIQE
Without SC-Attention	32.11	8.28	28.15
Without gater	32.67	8.33	28.57
Without loss (Only MSE loss)	33.21	8.47	29.43
Without dense feature fusion	32.95	8.53	28.89
URA-Net	31.35	7.98	27.77

Figs. 12 and 13 show the comparison of the defogging effect under each group of experiments. The SC attention mechanism, leveraging its sophisticated weighting mechanism, directs attention toward pivotal features such as dense foggy regions, high-frequency textures, and color fidelity. This targeted attention yields augmented image-defogging capabilities. The experiments conducted in this subsection adeptly demonstrate the constructive optimization influence stemming from the integration of the SC attention and self-attention gating mechanisms, alongside the utilization of the combined loss function within the network architecture. By assigning distinct weights for fusion operations to channel information characterized by varying receptive fields, the network garners enhanced adaptability, thereby fortifying its overall performance.

(a)   (b)   (c)   (d)   (e)   (f)   (g)

Figure 12: Ablation experimental qualitative evaluations (a) input image (b) without SC-Attention (c) without gater (d) without loss (e) without dense feature fusion (f) URA-net (g) groud truth

(a)   (b)   (c)   (d)   (e)   (f)

Figure 13: Ablation experimental qualitative evaluations on the real-world hazy image (a) input image (b) without SC-Attention (c) without gater (d) without loss (e) without dense feature fusion (f) URA-net

6 Experiments on Object Detection Task

Haze as an adverse weather condition, poses a significant challenge to the efficacy of land-based observation systems in remote sensing applications. In scenarios such as autonomous navigation, the presence of haze can severely impact object detection, leading to degraded image quality and potential safety risks. Therefore, it becomes imperative to employ a preprocessing step for image enhancement before undertaking tasks like object detection [126–130].

To address the absence of datasets containing built-in synthetic haze images for object detection, we conducted experiments using 100 randomly selected images from the VisDrone2019 datasets. Synthetic fogged versions of these images were generated for experimentation. For our model evaluations, we utilized YOLOv5 for target detection, with the dataset employing default outdoor pre-training weights. The experimental outcomes, as shown in Fig. 14, demonstrate improved detection performance post-defogging, underscoring the effectiveness of our proposed model in enhancing images for subsequent upstream tasks. This evidence supports the contention that our model contributes positively to addressing the challenges posed by haze in remote sensing applications, particularly in the context of object detection in adverse weather conditions.
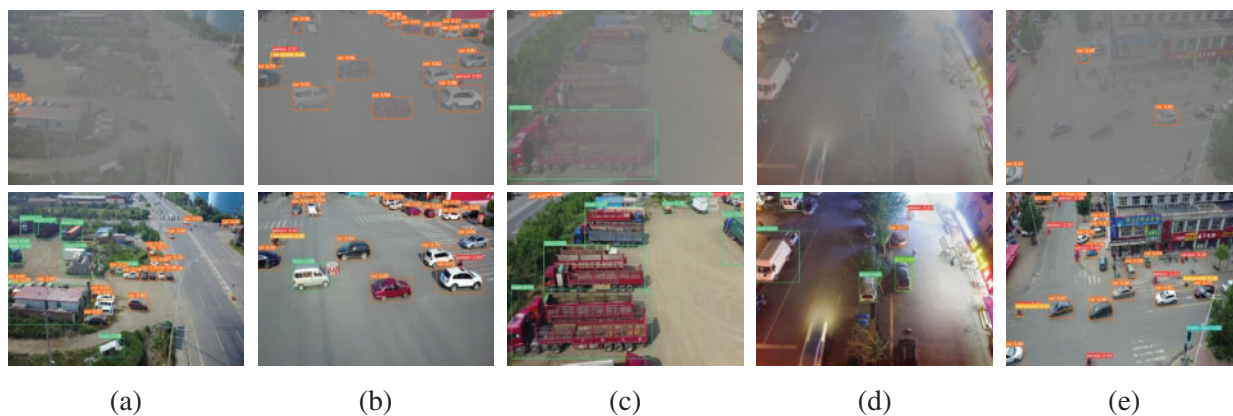

Figure 14: Comparison of object detection results under different conditions from VisDrone2019

Moreover, the detailed analysis of comprehensive detection results for a specific scene is visually presented in Fig. 14a, providing an overview of the overall impact of the de-fogging process. However, the crux of the improvement is vividly illustrated in Figs. 14b–14e, where the most compelling instances of the de-fogging effect are showcased. Noteworthy examples include Figs. 14b and 14c, which reveal a remarkable enhancement in the detection rate post-defogging. Here, additional instances of cars are successfully identified compared to the fogged condition, signifying the algorithm's efficacy in improving object recognition. In Fig. 14d, the de-fogging method addresses challenges posed by nighttime foggy images, effectively resolving detection issues and underscoring its capacity to enhance the accuracy of object recognition in challenging environments. Fig. 14e provides a before-and-after comparison, highlighting the transformative impact of the de-fogging process on a synthetic fogged image of a car. These visual examples not only validate the methodology's success in mitigating fog-induced obstacles but also emphasize its pivotal role in bolstering the reliability and accuracy of autonomous navigation systems operating under adverse weather conditions.

7 Limitations and Future Work

The URA-Net proposed in this paper has some limitations, which are shown in Fig. 15. (a) There is extremely dense and unevenly distributed haze in the test sample, which will lead to some irretrievable loss of image information and render the URA-Net algorithm powerless for hazy image restoration. (b) The proposed method will rely on a large number of labeled images for training, and it is difficult to achieve a good de-fogging effect for real foggy remote sensing images.
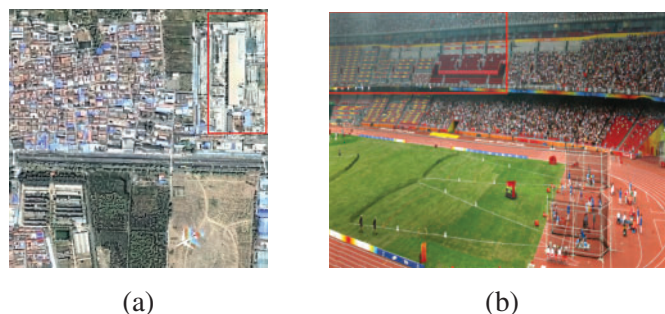
(a) (b)

Figure 15: Limitations of the URA-Net algorithm. (a) Distributed haze. (b) Real-world hazy image

The principal factor contributing to these limitations stems from the dependency on a substantial volume of labeled data for model training. However, these constraints can be mitigated through strategies such as domain adaptation and unsupervised learning. In our forthcoming research endeavors, we intend to delve deeper into the pertinent features associated with haze, refine our haze-adding algorithm, and explore the efficacy of deep learning networks in addressing the residual haze challenge inherent in real remote sensing image dehazing.

8 Conclusions

This paper presents a deep learning-based model tailored for remote sensing image defogging, encompassing a comprehensive review of image defogging methodologies. The initial section offers an overview of traditional techniques including image enhancement-based and physical model-based defogging, alongside deep learning methodologies. A meticulous literature comparison of deep learning-centric approaches is provided, elucidating the strengths and weaknesses of diverse strategies. The enhancement strategy adopted in this study is delineated thereafter. Subsequently, the paper delves into a detailed exposition of the proposed dense residual swarm-based defogging attention network. The discussion elucidates the intricacies of the model architecture and its operational mechanisms. Finally, the paper concludes with an assessment of the efficacy and superiority of the proposed model's defogging efficacy. This evaluation entails a comparative analysis of its performance against mainstream models leveraging the RICE-I and SateHaze1k datasets. The authors assert that the methodology introduced herein holds promise for elevating the quality of input images pertinent to foggy conditions. Consequently, this advancement is anticipated to bolster the accuracy and reliability of downstream tasks encompassing target detection, scene recognition, and image classification.

Author Contributions: The authors affirm their contributions to the paper as follows: Hongchi Liu, Xing Deng, and Haijian Shao were involved in the study conception, design, data collection, analysis, and interpretation of results, as well as in the preparation of the draft manuscript. All authors critically reviewed the results and provided their approval for the final version of the manuscript.

Availability of Data and Materials: The datasets utilized in the experiments include the RICE-I dataset [114] and the SateHaze1k dataset [115].

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Li Y, You S, Brown MS, Tan RT. Haze visibility enhancement: a survey and quantitative benchmarking. Comput Vis Image Underst. 2017;165:1–16. doi:10.1016/j.cviu.2017.09.003.

2. Hautière N, Tarel JP, Aubert D. Towards fog-free in-vehicle vision systems through contrast restoration. In: 2007 IEEE Conference on Computer Vision and Pattern Recognition; 2007; France. doi:10.1109/TIP.83.

3. Aggarwal S. Principles of remote sensing. Satellite remote sensing and GIS applications in agricultural meteorology. 2004;23(2):23–8.

4. Yadav G, Maheshwari S, Agarwal A. Fog removal techniques from images: a comparative review and future directions. In: 2014 International Conference on Signal Propagation and Computer Technology (ICSPCT 2014); 2014; Ajmer, India; p. 44–52.

5. Ju M, Ding C, Ren W, Yang Y, Zhang D, Guo YJ. IDE: image dehazing and exposure using an enhanced atmospheric scattering model. IEEE Trans Image Process. 2021;30:2180–92. doi:10.1109/TIP.83.

6. Song X, Zhou D, Li W, Dai Y, Shen Z, Zhang L, et al. TUSR-Net: triple unfolding single image dehazing with self-regularization and dual feature to pixel attention. IEEE Trans Image Process. 2023;32:1231–44. doi:10.1109/TIP.2023.3234701.

7. Zhang W, Dong L, Zhang T, Xu W. Enhancing underwater image via color correction and bi-interval contrast enhancement. Signal Process: Image Commun. 2021;90:116030.

8. Lei L, Wang L, Wu J, Bai X, Lv P, Wei M. Research on image defogging enhancement technology based on retinex algorithm. In: 2023 2nd International Conference on 3D Immersion, Interaction and Multi-sensory Experiences (ICDIIME); 2023; Madrid, Spain; p. 509–13.

9. Tang Q, Yang J, He X, Jia W, Zhang Q, Liu H. Nighttime image dehazing based on Retinex and dark channel prior using Taylor series expansion. Comput Vis Image Underst. 2021;202:103086. doi:10.1016/j.cviu.2020.103086.

10. Ni C, Fam PS, Marsani MF. Traffic image haze removal based on optimized retinex model and dark channel prior. J Intell Fuzzy Syst. 2022;43(6):8137–49. doi:10.3233/JIFS-221240.

11. Unnikrishnan H, Azad RB. Non-local retinex based dehazing and low light enhancement of images. Trait Sig. 2022;39(3):879–92.

12. Yu H, Li C, Liu Z, Guo Y, Xie Z, Zhou S. A novel nighttime dehazing model integrating Retinex algorithm and atmospheric scattering model. In: 2022 3rd International Conference on Geology, Mapping and Remote Sensing (ICGMRS); 2022; Zhoushan, China; p. 111–5, 2022.

13. Mondal K, Rabidas R, Dasgupta R. Single image haze removal using contrast limited adaptive histogram equalization based multiscale fusion technique. Multimed Tools Appl, 2024;83(5):15413–38.

14. Dharejo FA, Zhou Y, Deeba F, Jatoi MA, Du Y, Wang X. A remote-sensing image enhancement algorithm based on patch-wise dark channel prior and histogram equalisation with colour correction. IET Image Process. 2021;15(1):47–56. doi:10.1049/ipr2.v15.1.

15. Huang S, Zhang Y, Zhang O. Image haze removal method based on histogram gradient feature guidance. Int J Environ Res Public Health. 2023;20(4):3030. doi:10.3390/ijerph20043030.

16. Zhu Z, Hu J, Jiang J, Zhang X. A hazy image restoration algorithm via JND based histogram equalization and weighted DCP transmission factor. J Phy: Conf Series. 2021;1738(1):012035.

17. Abbasi N, Khan MF, Khan E, Alruzaiqi A, Al-Hmouz R. Fuzzy histogram equalization of hazy images: a concept using a type-2-guided type-1 fuzzy membership function. Granul Comput. 2023;8(4):731–45. doi:10.1007/s41066-022-00351-0.

18. Arif M, Badshah N, Khan TA, Ullah A, Rabbani H, Atta H, et al. A new Gaussian curvature of the image surface based variational model for haze or fog removal. PLoS One. 2023;18(3):e0282568. doi:10.1371/journal.pone.0282568.

19. Liu Y, Yan Z, Wu A, Ye T, Li Y. Nighttime image dehazing based on variational decomposition model. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops; 2022; New Orleans, LA, USA; p. 640–9.

20. Yu Z, Sun B, Liu D, De Dravo VW, Khokhlova M, Wu S. STRASS dehazing: spatio-temporal retinex-inspired dehazing by an averaging of stochastic samples. J Renew Mater. 2022;10(5):1381–1395. doi:10.32604/jrm.2022.018262.

21. Dai W, Ren X. Defogging algorithm for road environment landscape visual image based on wavelet transform. In: 2023 International Conference on Networking, Informatics and Computing (ICNETIC); 2023; Palermo, Italy; p. 587–91.

22. Liang W, Long J, Li KC, Xu J, Ma N, Lei X. A fast defogging image recognition algorithm based on bilateral hybrid filtering. ACM Trans Multimed Comput, Commun Appl (TOMM). 2021;17(2):1–16.

23. Lai Y, Zhou Z, Su B, Xuanyuan Z, Tang J, Yan J, et al. Single underwater image enhancement based on differential attenuation compensation. Front Mar Sci. 2022;9:1047053. doi:10.3389/fmars.2022.1047053.

24. Chen Z, Wang Y, Yang Y, Liu D. PSD: principled synthetic-to-real dehazing guided by physical priors. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2021; Nashville, TN, USA; p. 7180–9.

25. Zhou H, Chen Z, Liu Y, Sheng Y, Ren W, Xiong H. Physical-priors-guided DehazeFormer. Knowl-Based Syst. 2023;266:110410. doi:10.1016/j.knosys.2023.110410

26. Gupta R, Khari M, Gupta V, Verdú E, Wu X, Herrera-Viedma E, et al. Fast single image haze removal method for inhomogeneous environment using variable scattering coefficient. Comput Model Eng Sci. 2020;123(3):1175–92. doi:10.32604/cmes.2020.010092.

27. Wu RQ, Duan ZP, Guo CL, Chai Z, Li C. RIDCP: revitalizing real image dehazing via high-quality codebook priors. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2023; Vancouver, BC, Canada; p. 22282–91.

28. Zhang ZP, Cheng Y, Zhang S, Bu LJ, Deng MJ. A multimodal feature fusion image dehazing method with scene depth prior. IET Image Process. 2023;17(11):3079–94. doi:10.1049/ipr2.v17.11.

29. He K, Sun J, Tang X. Single image haze removal using dark channel prior. IEEE Trans Pattern Anal Mach Intell. 2010;33(12):2341–53.

30. Pan Y, Chen Z, Li X, He W. Single-image dehazing via dark channel prior and adaptive threshold. Int J Image Graph. 2021;21(04):2150053. doi:10.1142/S0219467821500534.

31. Xu Y. IOP publishing. an adaptive dehazing network based on dark channel prior and image segmentation. J Phy: Conf Series. 2023;2646(1):012009.

32. Li C, Yuan C, Pan H, Yang Y, Wang Z, Zhou H, et al. Single-image dehazing based on improved bright channel prior and dark channel prior. Electronics. 2023;12(2):299. doi:10.3390/electronics12020299.

33. Yang G, Yang H, Yu S, Wang J, Nie Z. A multi-scale dehazing network with dark channel priors. Sensors. 2023;23(13):5980. doi:10.3390/s23135980.

34. Aksas L, Lapray PJ, Foulonneau A, Bigué L. Joint qualitative and quantitative evaluation of fast image dehazing based on dark channel prior. Unconvent Opt Imag III. 2022;12136:295–303.

35. He Y, Li C, Li X. Remote sensing image dehazing using heterogeneous atmospheric light prior. IEEE Access. 2023;11:18805–20. doi:10.1109/ACCESS.2023.3247967.

36. Xu Y, Zhang H, He F, Guo J, Wang Z. Enhanced CycleGAN network with adaptive dark channel prior for unpaired single-image dehazing. Entropy. 2023;25(6):856. doi:10.3390/e25060856.

37. Yu J, Zhang J, Li B, Ni X, Mei J. Nighttime image dehazing based on bright and dark channel prior and gaussian mixture model. In: Proceedings of the 2023 6th International Conference on Image and Graphics Processing; 2023; Chongqing, China; p. 44–50.

38. Hu Q, Zhang Y, Liu T, Liu J, Luo H. Maritime video defogging based on spatial-temporal information fusion and an improved dark channel prior. Multimed Tools Appl. 2022;81(17):24777–98. doi:10.1007/s11042-022-11921-4.

39. Thomas J, Raj ED. Single image dehazing with color correction transform dark channel prior. Social Science Electronic Publishing; 2022.

40. Chu Y, Chen F, Fu H, Yu H. Haze level evaluation using dark and bright channel prior information. Atmosphere 2022;13(5):683. doi:10.3390/atmos13050683.

41. Mo Y, Li C, Zheng Y, Wu X. DCA-CycleGAN: unsupervised single image dehazing using dark channel attention optimized CycleGAN. J Vis Commun Image Represent. 2022;82:103431. doi:10.1016/j.jvcir.2021.103431.

42. Miao L, Yan L, Jin Z, Yuqian Z, Shuanhu D. Image dehazing method based on haze-line and color attenuation prior. J Commun/Tongxin Xuebao. 2023:44(1):211–22.

43. Li ZX, Wang YL, Peng C, Peng Y. Laplace dark channel attenuation-based single image defogging in ocean scenes. Multimed Tools Appl. 2023;82(14):21535–59. doi:10.1007/s11042-022-14103-4.

44. Soma P, Jatoth RK. An efficient and contrast-enhanced video de-hazing based on transmission estimation using HSL color model. Vis Comput. 2022;38(7):2569–80. doi:10.1007/s00371-021-02132-3.

45. Nandini B, Kaulgud N. Wavelet-based method for enhancing the visibility of hazy images using Color Attenuation Prior. In: 2023 International Conference on Recent Trends in Electronics and Communication (ICRTEC); 2023; Mysore, p. 1–6.

46. Krichen M. Convolutional neural networks: a survey. Computers. 2023;12(8):151. doi:10.3390/computers12080151.

47. Sarvamangala D, Kulkarni RV. Convolutional neural networks in medical image understanding: a survey. Evol Intell. 2022;15(1):1–22. doi:10.1007/s12065-020-00540-3.

48. Zhang X, Zhang X, Wang W. Convolutional neural network. Intell Inform Process with Matlab. Germany; 2023; p. 39–71.

49. Zafar A, Aamir M, Mohd Nawi N, Arshad A, Riaz S, Alruban A, et al. A comparison of pooling methods for convolutional neural networks. Appl Sci. 2022;12(17):8643. doi:10.3390/app12178643.

50. Mehak, Whig P. More on convolution neural network CNN. International Journal of Sustainable Development in Computing Science. 2022;4(1).

51. Pomazan, V, Tvoroshenko, I, Gorokhovatskyi, V. Development of an application for recognizing emotions using convolutional neural networks. Int J Acad Inform Syst Res. 2023;7(7):25–36.

52. Riad R, Teboul O, Grangier D, Zeghidour N. Learning strides in convolutional neural networks. arXiv preprint arXiv:220201653. 2022.

53. Kumar, M. S, Ganesh, D, Turukmane, A. V, Batta, U, Sayyadliyakat, K. K.. Deep convolution neural network based solution for detecting plant diseases. J Pharm Negat Results. 2022;13, 464–71.

54. Cai B, Xu X, Jia K, Qing C, Tao D. DehazeNet: an end-to-end system for single image haze removal. IEEE Trans Image Process. 2016;25(11):5187–98. doi:10.1109/TIP.2016.2598681.

55. Wang C, Zou Y, Chen Z. ABC-NET: avoiding blocking effect & color shift network for single image dehazing via restraining transmission bias. In: 2020 IEEE International Conference on Image Processing (ICIP); 2020; Abu Dhabi, United Arab Emirates; p. 1053–7.

56. Zhu Z, Luo Y, Wei H, Li Y, Qi G, Mazur N, et al. Atmospheric light estimation based remote sensing image dehazing. Remote Sens. 2021;13(13):2432. doi:10.3390/rs13132432.

57. Liu S, Li Y, Li H, Wang B, Wu Y, Zhang Z. Visual image dehazing using polarimetric atmospheric light estimation. Appl Sci. 2023;13(19):10909. doi:10.3390/app131910909.

58. Hong S, Kang MG. Single image dehazing based on pixel-wise transmission estimation with estimated radiance patches. Neurocomputing. 2022;492:545–60. doi:10.1016/j.neucom.2021.12.046.

59. Yang S, Sun Z, Jiang Q, Zhang Y, Bao F, Liu P. A mixed transmission estimation iterative method for single image dehazing. IEEE Access. 2021;9:63685–99. doi:10.1109/ACCESS.2021.3074531.

60. Muniraj M, Dhandapani V. Underwater image enhancement by combining color constancy and dehazing based on depth estimation. Neurocomputing. 2021;460:211–30. doi:10.1016/j.neucom.2021.07.003.

61. Sun Z, Zhang Y, Bao F, Shao K, Liu X, Zhang C. ICycleGAN: single image dehazing based on iterative dehazing model and CycleGAN. Comput Vis Image Underst. 2021;203:103133. doi:10.1016/j.cviu.2020.103133.

62. Wang C, Huang Y, Zou Y, Xu Y. Fully non-homogeneous atmospheric scattering modeling with convolutional neural networks for single image dehazing. arXiv preprint arXiv:210811292. 2021.

63. Wang C, Huang Y, Zou Y, Xu Y. FWB-Net: front white balance network for color shift correction in single image dehazing via atmospheric light estimation. In: ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2021; Toronto, ON, Canada; p. 2040–44.

64. Song Y, He Z, Qian H, Du X. Vision transformers for single image dehazing. IEEE Trans Image Process. 2023;32:1927–41. doi:10.1109/TIP.2023.3256763.

65. Bai H, Pan J, Xiang X, Tang J. Self-guided image dehazing using progressive feature fusion. IEEE Trans Image Process. 2022;31:1217–29. doi:10.1109/TIP.2022.3140609.

66. Wang S, Chen P, Huang J, Wong TH. FPD Net: feature pyramid dehazenet. Comput Syst Sci Eng. 2022;40(3):1167–81. doi:10.32604/csse.2022.018911.

67. Riaz S, Anwar MW, Riaz I, Kim HW, Nam Y, Khan MA. Multiscale image dehazing and restoration: an application for visual surveillance. Comput Mater Contin. 2021;70:1–17. doi:10.32604/cmc.2022.018268.

68. Ullah H, Muhammad K, Irfan M, Anwar S, Sajjad M, Imran AS, et al. Light-DehazeNet: a novel lightweight CNN architecture for single image dehazing. IEEE Trans Image Process. 2021;30:8968–82. doi:10.1109/TIP.2021.3116790.

69. Feng T, Wang C, Chen X, Fan H, Zeng K, Li Z. URNet: a U-Net based residual network for image dehazing. Appl Soft Comput. 2021;102:106884. doi:10.1016/j.asoc.2020.106884.

70. Huang P, Zhao L, Jiang R, Wang T, Zhang X. Self-filtering image dehazing with self-supporting module. Neurocomputing. 2021;432:57–69. doi:10.1016/j.neucom.2020.11.039.

71. Zhang X, Wang J, Wang T, Jiang R. Hierarchical feature fusion with mixed convolution attention for single image dehazing. IEEE Trans Circuits Syst Video Technol. 2021;32(2):510–22.

72. Ali A, Ghosh A, Chaudhuri SS. LIDN: a novel light invariant image dehazing network. Eng Appl Artif Intell. 2023;126:106830. doi:10.1016/j.engappai.2023.106830.

73. Chen D, He M, Fan Q, Liao J, Zhang L, Hou D, et al. Gated context aggregation network for image dehazing and deraining. In: 2019 IEEE Winter Conference on Applications of Computer Vision (WACV); 2019; Waikoloa, HI, USA; p. 1375–83.

74. Ren W, Ma L, Zhang J, Pan J, Cao X, Liu W, et al. Gated fusion network for single image dehazing. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA; p. 3253–61.

75. Liu X, Ma Y, Shi Z, Chen J. GridDehazeNet: attention-based multi-scale network for image dehazing. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV); 2019; Korea (South); p. 7314–23.

76. Qin X, Wang Z, Bai Y, Xie X, Jia H. FFA-Net: feature fusion attention network for single image dehazing. In: Proceedings of the AAAI Conference on Artificial Intelligence; 2020; New York, USA; p. 11908–15.

77. Zhang J, Cao Y, Zha ZJ, Tao D. Nighttime dehazing with a synthetic benchmark. In: Proceedings of the 28th ACM International Conference on Multimedia; 2020; Seattle, USA; p. 2355–63.

78. Chen, E, Chen, S, Ye, T, Liu, Y. Degradation-adaptive neural network for jointly single image dehazing and desnowing. Front Comput Sci. 2024;18(2):1–3.

79. Yang X, Xu Z, Luo J. Towards perceptual image dehazing by physics-based disentanglement and adversarial training. In: Proceedings of the AAAI Conference on Artificial Intelligence; 2018; Louisiana, USA. doi:10.1609/aaai.v32i1.12317.

80. Sun Z, Zhang Y, Bao F, Shao K, Liu X, Zhang C. ICycleGAN: single image dehazing based on iterative dehazing model and CycleGAN. Comput Vis Image Underst. 2021;203:103133. doi:10.1016/j.cviu.2020.103133.

81. Shyam P, Yoo H. Data efficient single image dehazing via adversarial auto-augmentation and extended atmospheric scattering model. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops; 2023; Paris, France; p. 227–37.

82. Lin S, Sun P, Gao H, Ju Z. Haze optical-model-based nighttime image dehazing by modifying attenuation and atmospheric light. JOSA A. 2022;39(10):1893–902. doi:10.1364/JOSAA.463033.

83. Bhavani BD, Rohini A, Sravika A, Harini A. Enhancing atmospheric visibility in satellite images using dehazing methods. J Eng Sci. 2023;14(10):1–13.

84. Boob D, Dey SS, Lan G. Complexity of training relu neural network. Discrete Optim. 2022;44:100620. doi:10.1016/j.disopt.2020.100620.

85. Gupta M, Camci E, Keneta VR, Vaidyanathan A, Kanodia R, Foo CS, et al. Is complexity required for neural network pruning? a case study on global magnitude pruning. arXiv preprint arXiv: 220914624. 2022.

86. Lu Y, Chen Y, Zhao D, Chen J. Graph-FCN for image semantic segmentation. In: Advances in Neural Networks–ISNN 2019. Moscow, Russia; 2019; p. 97–105.

87. Adegun AA, Viriri S. FCN-based DenseNet framework for automated detection and classification of skin lesions in dermoscopy images. IEEE Access. 2020;8:150377–96. doi:10.1109/Access.6287639.

88. Jian M, Wu R, Chen H, Fu L, Yang C (2023). Dual-branch-UNet: a dual-branch convolutional neural network for medical image segmentation. Comput Model Eng Sci. 137(1):705–16. doi:10.32604/cmes.2023.027425.

89. Yu L, Qin Z, Ding Y, Qin Z. MIA-UNet: multi-scale iterative aggregation u-network for retinal vessel segmentation. Comput Model Eng Sci. 2021;129(2):805–28. doi:10.32604/cmes.2021.017332.

90. Bapatla S, Harikiran J. LuNet-LightGBM: an effective hybrid approach for lesion segmentation and DR grading. Comput Syst Sci Eng. 2023;46(1):597–617. doi:10.32604/csse.2023.034998.

91. Kwon YM, Bae S, Chung DK, Lim MJ. Semantic segmentation by using down-sampling and subpixel convolution: dSSC-UNet. Comput Mater Contin. 2023;75(1):683–96. doi:10.32604/cmc.2023.033370.

92. Hayou S, Clerico E, He B, Deligiannidis G, Doucet A, Rousseau J. PMLR. Stable resnet. Int Conf Artif Intell Stat, 2021; p. 1324–32.

93. Zhang C, Benz P, Argaw DM, Lee S, Kim J, Rameau F, et al. Resnet or densenet? Introducing dense shortcuts to resnet. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV); 2021; Waikoloa, HI, USA; p. 3550–9.

94. Tran ST, Cheng CH, Nguyen TT, Le MH, Liu DG. TMD-unet: triple-unet with multi-scale input features and dense skip connection for medical image segmentation. Healthcare. 2021;9(1):54. doi:10.3390/healthcare9010054.

95. Qian L, Wen C, Li Y, Hu Z, Zhou X, Xia X, et al. Multi-scale context UNet-like network with redesigned skip connections for medical image segmentation. Comput Methods Programs Biomed. 2024;243:107885. doi:10.1016/j.cmpb.2023.107885.

96. Oommen DK, Arunnehru J. Alzheimer's disease stage classification using a deep transfer learning and sparse auto encoder Method. Comput Mater Contin. 2023;76(1):793–811. doi:10.32604/cmc.2023.038640.

97. Zhao L, Zhang Y, Cui Y. An attention encoder-decoder network based on generative adversarial network for remote sensing image dehazing. IEEE Sens J. 2022;22(11):10890–900. doi:10.1109/JSEN.2022.3172132.

98. Niu Z, Zhong G, Yu H. A review on the attention mechanism of deep learning. Neurocomputing. 2021;452:48–62. doi:10.1016/j.neucom.2021.03.091.

99. Shu X, Zhang L, Qi GJ, Liu W, Tang J. Spatiotemporal co-attention recurrent neural networks for human-skeleton motion prediction. IEEE Trans Pattern Anal Mach Intell. 2021;44(6):3300–15.

100. Mittal S, Raparthy SC, Rish I, Bengio Y, Lajoie G. Compositional attention: disentangling search and retrieval. arXiv preprint arXiv:211009419. 2021.

101. Niu Z, Zhong G, Yu H. A review on the attention mechanism of deep learning. Neurocomputing. 2021;452:48–62. doi:10.1016/j.neucom.2021.03.091.

102. Zhu Y, Newsam S. DenseNet for dense flow. In: 2017 IEEE International Conference on Image Processing (ICIP); 2017; Beijing, China; p. 790–4.

103. Dong H, Pan J, Xiang L, Hu Z, Zhang X, Wang F, et al. Multi-scale boosted dehazing network with dense feature fusion. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2020; Seattle, WA, USA; p. 2157–67.

104. Shan L, Wang W. DenseNet-based land cover classification network with deep fusion. IEEE Geosci Remote Sens Lett. 2021;19:1–5.

105. Liu X, Chen S, Song L, Woźniak M, Liu S. Self-attention negative feedback network for real-time image super-resolution. J King Saud Univ-Comput. Inf Sci. 2022;34(8):6179–86.

106. Ding J, Guo H, Zhou H, Yu J, He X, Jiang B. Distributed feedback network for single-image deraining. Inf Sci. 2021;572:611–26. doi:10.1016/j.ins.2021.02.080.

107. Liu L, Song X, Lyu X, Diao J, Wang M, Liu Y, et al. FCFR-Net: Feature fusion based coarse-to-fine residual learning for depth completion. In: Proceedings of the AAAI Conference on Artificial Intelligence; 2021; Vancouver, Canada; p. 2136–44.

108. Gurrola-Ramos J, Dalmau O, Alarcón TE. A residual dense u-Net neural network for image denoising. IEEE Access. 2021;9:31742–54. doi:10.1109/Access.6287639.

109. Li F, Luo M, Hu M, Wang G, Chen Y. Liver tumor segmentation based onmulti-scale and self-attention mechanism. Comput Syst Sci Eng. 2023;47(3):2835–50. doi:10.32604/csse.2023.039765.

110. Tong X, Wei J, Sun B, Su S, Zuo Z, Wu P. ASCU-Net: attention gate, spatial and channel attention u-net for skin lesion segmentation. Diagnostics. 2021;11(3):501. doi:10.3390/diagnostics11030501.

111. Hodson TO. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. Geosci Model Dev. 2022;15(14):5481–7. doi:10.5194/gmd-15-5481-2022.

112. Setiadi DRIM. PSNR vs SSIM: imperceptibility quality assessment for image steganography. Multimed Tools Appl. 2021;80(6):8423–44. doi:10.1007/s11042-020-10035-z.

113. Bakurov I, Buzzelli M, Schettini R, Castelli M, Vanneschi L. Structural similarity index (SSIM) revisited: a data-driven approach. Expert Syst Appl. 2022;189:116087. doi:10.1016/j.eswa.2021.116087.

114. Lin D, Xu G, Wang X, Wang Y, Sun X, Fu K. A remote sensing image dataset for cloud removal. arXiv preprint arXiv:190100600. 2019.

115. Huang B, Zhi L, Yang C, Sun F, Song Y. Single satellite optical imagery dehazing using SAR image prior based on conditional generative adversarial networks. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV); 2020; Snowmass, CO, USA; p. 1806–13.

116. He Y, Li C, Li X. Remote sensing image dehazing using heterogeneous atmospheric light prior. IEEE Access. 2023;11:18805–20. doi:10.1109/ACCESS.2023.3247967.

117. Li B, Peng X, Wang Z, Xu J, Feng D. AOD-Net: All-in-One dehazing network. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV); 2017; Italy; p. 4770–8.

118. Qin X, Wang Z, Bai Y, Xie X, Jia H. FFA-Net: Feature fusion attention network for single image dehazing. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2020; New York, USA; p. 11908–15.

119. Chen D, He M, Fan Q, Liao J, Zhang L, Hou D, et al. Gated context aggregation network for image dehazing and deraining. In: 2019 IEEE Winter Conference on Applications of Computer Vision (WACV); 2019; Waikoloa, HI, USA; p. 1375–83.

120. Ngo D, Lee S, Kang B. Robust single-image haze removal using optimal transmission map and adaptive atmospheric light. Remote Sens. 2020;12(14):2233. doi:10.3390/rs12142233.

121. Li J, Li Y, Zhuo L, Kuang L, Yu, T. Unsupervised single image dehazing network via disentangled representations. IEEE Trans Multimed. 2023;25:3587–601.

122. He Y, Li C, Bai T. Remote sensing image haze removal based on superpixel. Remote Sens. 2023;15(19):4680. doi:10.3390/rs15194680.

123. Mittal A, Moorthy AK, Bovik AC. No-reference image quality assessment in the spatial domain. IEEE Trans Image Process. 2012;21(12):4695–708. doi:10.1109/TIP.2012.2214050.

124. Mittal A, Soundararajan R, Bovik AC. Making a "completely blind" image quality analyzer. IEEE Signal Process Lett. 2012;20(3):209–12.

125. Venkatanath N, Praneeth D, Bh MC, Channappayya SS, Medasani SS. Blind image quality evaluation using perception based features. In: 2015 Twenty First National Conference on Communications (NCC); 2015; Mumbai, India; p. 1–6.

126. Liu W, Ren G, Yu R, Guo S, Zhu J, Zhang L. Image-adaptive YOLO for object detection in adverse weather conditions. In: Proceedings of the AAAI Conference on Artificial Intelligence; 2022; BC, Canada; p. 1792–800.

127. Wang H, Xu Y, He Y, Cai Y, Chen L, Li Y, et al. YOLOv5-Fog: a multiobjective visual detection algorithm for fog driving scenes based on improved YOLOv5. IEEE Trans Instrum Meas. 2022;71:1–12.

128. Zhang Q, Hu X. MSFFA-YOLO network: multi-class object detection for traffic investigations in foggy weather. IEEE Trans Instrum Meas. 2023;72:2528712.

129. Zhang Y, Ge H, Lin Q, Zhang M, Sun Q. Research of maritime object detection method in foggy environment based on improved model SRC-YOLO. Sensors. 2022;22(20):7786. doi:10.3390/s22207786.

130. Huang J, He Z, Guan Y, Zhang H. Real-time forest fire detection by ensemble lightweight YOLOX-L and defogging method. Sensors. 2023;23(4):1894. doi:10.3390/s23041894.