



OPEN Runge-kutta method inspired aerial image dehazing network in YUV space

Yazhong Si¹, Jingyu Chen^{1,3}, Wenqing Rao^{1,3} & Chen Li²✉

In hazy weather, the quality of aerial images suffers severe degradation, which affects the imaging capabilities of advanced remote sensing applications. Most existing learning-based dehazing algorithms utilize manually designed deep structures to enhance model performance. However, these deep networks often contain redundant branches, leading to a significant decrease in computational efficiency. To address this issue, we first conduct experiments to explore the distribution of haze and introduce an improved dehazing strategy in the YUV space. Subsequently, we propose a Runge-Kutta (RK) method-inspired aerial image dehazing network called RKNet, which consists of two parts: luminance (Y) domain haze removal and chrominance (U,V) enhancement. From the perspective of dynamical systems and based on the 3rd-order RK method, we design an RK3 block and incorporate it into RKNet to improve computational accuracy. Experimental evaluations on both synthetic and real-world benchmarks demonstrate that RKNet outperforms current haze removal algorithms and achieves superior performance. In addition, RKNet can improve the detection accuracy and capability of high-level vision algorithms and is also applicable to the processing of sandstorm images and underwater images.

Keywords Aerial image, Image dehazing, YUV space, Runge-kutta method

Hazy weather is a common atmospheric phenomenon that adversely effects on both human activities and machine vision systems. When captured in foggy conditions, images generally appear visually hazy and blurred. Such degraded images are unsuitable for applications that necessitate accurate environmental information for safe operation¹, such as unmanned aerial drones, autonomous vehicles, and intelligent infrastructure. As a result, there has been a growing interest in single-image haze removal as a critical low-level computer vision task in recent years. The atmospheric scattering model provides a simplified approximation for the haze imaging process, which is written as:

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

To restore the haze-free image, the atmospheric scattering model can be converted into the following form:

$$J(x) = A + \frac{I(x) - A}{t(x)} \quad (2)$$

$$t(x) = e^{-\beta d(x)} \quad (3)$$

where $I(x)$ and $J(x)$ denote the haze image and clear image respectively; $t(x)$ is the transmission, β is the scattering coefficient of the atmosphere, $d(x)$ is the scene depth; A represents the global atmosphere light value.

The main challenge in atmospheric scattering model-based haze restoration algorithms is how to estimate the transmission map $t(x)$ and atmospheric light A accurately. Existing approaches generally calculate the intermediate parameters via the prior knowledge²⁻⁴. However, these methods are limited as prior knowledge can not accurately describe the characteristics of the scenes, which will lead to inaccurate estimation of the parameters, resulting in the issues of blurring and color distortion in outputs.

Recent years have witnessed the significant success of convolutional neural networks (CNN) in high-level computer vision applications⁵⁻⁷. Building on the prior work, Cai et al.⁸ introduced a CNN-based dehazing

¹School of Electronics and Information Engineering, Wuxi University, 214105 Wuxi, China. ²School of Cybersecurity and Informatization, Wuxi University, 214105 Wuxi, China. ³School of Electronic and Information Engineering, Nanjing University of Information Science & Technology, 210044 Nanjing, China. ✉email: Papertg2018@163.com

algorithm. Since then, CNN has become the predominant method in the field of image haze removal^{9–11}. While efforts have been made to improve performance, there are still some limitations in current dehazing methods. Firstly, many studies prioritize performance over computation overhead, leading to the algorithms that may be too computationally intensive for practical deployment. Secondly, training deeper networks requires more advanced numerical instability avoidance techniques. Thirdly, unlike high-level visual tasks such as image segmentation and object detection that extract semantic features through CNN, image dehazing requires prediction of fine-grained pixel-level details, making direct use of state-of-the-art (SOTA) CNN suboptimal.

Actually, compared with RGB space, YUV space can provide a more accurate description of haze degradation characteristics¹². To address the aforementioned issues, we first analyze the haze distribution and explore the feasibility of the dehazing strategy in YUV space. Then, we integrate the Runge-Kutta (RK) method-inspired scheme into haze removal network designs and name the model as RKNNet. Experimental results on benchmarks demonstrate that RKNNet outperforms existing SOTA methods, achieving a superior balance between computing cost and performance. Figure 1 presents the dehazing results of RKNNet on LHID dataset¹³, and it can be seen that compared with the original haze image, the reconstructed image has a significant improvement in visual clarity. To our best knowledge, this is the first attempt to directly incorporate an RK method-inspired scheme into the design of image dehazing network.

The main contributions of this study are summarized as follows:

- We explore the distribution characteristics of haze in YUV space and propose a joint processing method for Y-channel dehazing and UV-channel enhancement.
- Inspired by the Runge-Kutta method, we introduce a novel aerial image dehazing network named RKNNet, which achieves a remarkable balance between performance and computational cost.
- Extensive experiments are performed on both synthetic and real-world haze images, demonstrating the superiority of the proposed RKNNet over the current SOTAs.

Related work

Image haze removal

Single image haze removal as a classical computer vision task has garnered significant interest from researchers. Existing dehazing algorithms can be broadly categorized into two groups: prior knowledge-driven methods and CNN-based methods.

Prior knowledge-driven methods

The prior knowledge-driven methods are the reverse procedure of the atmospheric scattering model described as Eq. (2). They estimate the necessary parameters in atmospheric scattering model by using the prior assumptions

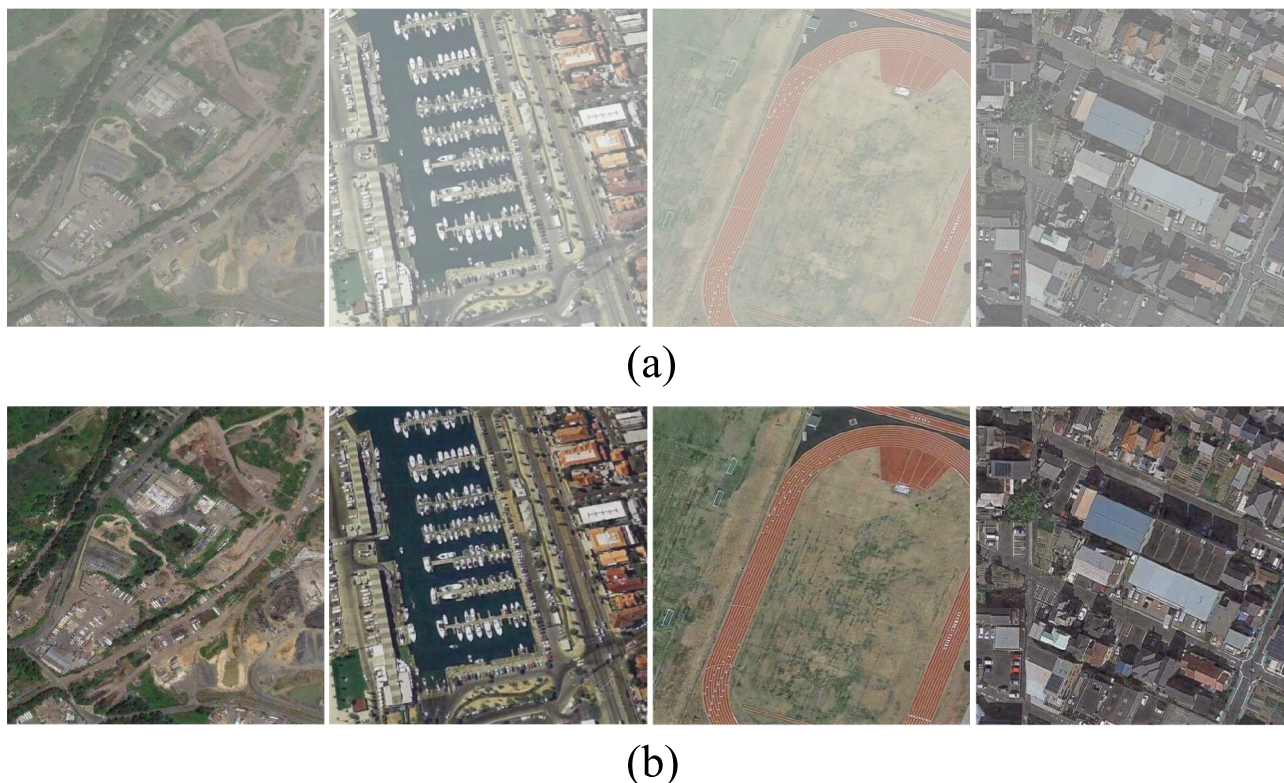


Fig. 1. The dehazing results of the proposed method. (a) Aerial haze images; (b) The enhanced results of RKNNet.

which are derived from the statistical characteristics of the haze images. He et al.² established the dark channel prior (DCP) theory, which is built on a statistical observation that there are some dark pixels close to zero in at least one color channel within the non-sky area of an outdoor haze-free image. Bui et al.³ introduced the color ellipsoid prior (CEP), where they construct color ellipsoids that are statistically fitted to haze pixel clusters in RGB space. They then calculate the transmission values based on the geometry of these color ellipsoids. Yadav et al.¹⁴ proposed a hazy image restoration method, which estimates the approximate depth map of hazy images to derive scene-specific dark channels and transmissions, incorporates histogram normalization for post-processing to enhance image quality. Kaur et al.⁴ proposed the gradient channel prior (GCP), which estimates the transmission map via the image gradient. Liu et al.¹⁵ presented a straightforward method for estimating the transmission map based on the rank-one prior (ROP). Following the proposed ROP theory, they¹⁶ took the foreground and the background into consideration, developing ROP⁺ to further improve the performance of the algorithm.

Although these algorithms can improve the clarity of haze images to some extent, there are still some issues worth considering. Prior knowledge-driven methods rely on fixed assumptions grounded in mathematical statistics, which may not be effective in some real-world scenarios. In addition, inaccurate parameter estimation will lead to artifacts, halos, and color distortion in the restored results.

CNN-based methods

Early CNN-based dehazing algorithms still rely on the atmospheric scattering model¹⁷⁻¹⁹. These methods employ CNN to estimate the transmission map and atmospheric light value from haze images, which are then used to inversely infer the model and obtain clear restored images. For instance, Cai et al.⁸ proposed a dehazing network embedded with bilateral rectified linear units to estimate the medium transmission map. Ren et al.²⁰ developed a coarse-to-fine multiscale deep neural network to predict the refined transmission map. Su et al.²¹ proposed a fusion network integrating a physical model and image translation for image dehazing, incorporating the estimation of the transmission map through a deep network to guide image translation. Zhang et al.²² introduced a pyramid densely connected transmission map estimation network based on generative adversarial networks (GAN). However, the transmission map is susceptible to noise interference, which will weaken the robustness of the algorithms.

To address the issue, researchers are starting to focus on end-to-end image dehazing methods that directly establish the mapping between haze images and clear images without relying on the physical model. Qin et al.²³ introduced a feature attention module (FA) that combines channel attention with pixel attention mechanism, embedding it into the dehazing network. Wu et al.⁹ proposed an autoencoder-like dehazing network inspired by contrastive learning. Wang et al.²⁴ introduced an unsupervised contrastive learning paradigm called UCL-Dehaze for image dehazing, which leverages unpaired real-world hazy and clean images through adversarial training and a new pixel-wise self-contrastive perceptual loss. Bai et al.¹¹ extracted feature information from the haze image itself to guide network training. Su et al.²⁵ proposed an end-to-end dehazing network with a parameter-shared architecture that trains on synthetic and real haze images simultaneously, using multi-prior pseudo clean image and a physical-model-guided domain transfer mechanism to reduce domain gaps. Although CNN-based haze removal methods have shown some performance success, the SOTA designs heavily rely on empirical structure, involving numerous attempts and tricks. Image enhancement as a foundational low-level vision task, hinges on recovering fine-grained details like edges, contours, and textures obscured by haze, and over-reliance on abstract high-level features undermines sharpness and clarity, making preservation of shallow low-level features critical for visually faithful results^{26,27}. We thus consider that simplifying the dehazing algorithm into a low-level work would be a beneficial approach.

ODE-inspired network design

With the advancement of CNN, scholars have started to explore the theoretical properties of CNN from the perspective of ordinary differential equations (ODE). Weinan et al.²⁸ were the first to observe the connection between residual network (ResNet)²⁹ and ODE. They demonstrated that deep neural networks can be treated as discrete dynamical systems, and established the similarities between ResNet and the discretization of ODE. Chang et al.³⁰ went beyond mere explanation and utilized numerical ODEs to construct reversible neural networks while conducting stability analysis. Inspired by the linear multi-step method for solving ODEs, Lu et al.³¹ proposed a linear multi-step architecture (LM-architecture). They integrated the LM-architecture into ResNet and ResNeXt³², and observed significantly improved accuracy compared to the original network on public benchmarks.

In the field of low-level image translation, recent studies³³⁻³⁶ have established a connection between ResNet and the explicit/implicit Euler method approximation of ODE. However, the research on ODE-based dehazing algorithms is still in its infancy, and there are numerous issues that warrant further exploration. Given that image haze removal is a low-level computer vision task with the constraint of the highly similar structure between haze and haze-free images, it often relies on intuitive approaches. Building on the inspiration from the aforementioned work, we introduce a simple yet effective image dehazing algorithm.

The proposed method

The progress in single image dehazing can be attributed to the advancements in deep learning, which provides a powerful framework for nonlinear calculations in haze removal algorithms. In general, CNN-based algorithms consist of multiple cascaded layers and can effectively transform complex information flow after iterative training. In a dynamical system, there exists a fixed rule that describes the evolution of input states in a geometric space over time. Similarly, the adaptively selected layers in a CNN can be treated as time nodes in a dynamical system, with the final output state constrained by the ground truth (GT). Previous work²⁸ has characterized this as a

controllable issue and conducted a simplified analysis of the low-dimensional version, concluding that a solution can be obtained using numerical methods if the issue is sufficiently smooth. Encouraged by this, we apply the Runge-Kutta (RK) method to our haze removal network design from the perspective of the dynamical system. In this section, we will first analyze the distribution of haze in YUV space and then discuss the details of RKNNet.

YUV space analysis and the improved strategy

In accordance with the prescribed guidelines outlined in the *ITU-R BT.601-7* standards established by the Radio Communication Division of the International Telecommunication Union (ITU-R), the conversion from RGB to YUV in full range³⁷ is outlined as follows:

$$Y = 0.299 * R + 0.587 * G + 0.114 * B \quad (4)$$

$$U = \frac{B - Y}{1.772} = -0.169 * R - 0.331 * G + 0.500 * B \quad (5)$$

$$V = \frac{R - Y}{1.402} = 0.500 * R - 0.419 * G - 0.081 * B \quad (6)$$

Recent research has proved that YUV space can provide a more effective way for image dehazing¹². Specifically, they separated the luminance (Y) and chrominance (UV) components and found that the energy of the haze is concentrated in Y channel, while UV channel is less affected. Figure 2 presents a comparison of indicators in RGB and YUV space for RS haze images from the SateHaze1k dataset. In the upper and lower rows, from left to right, are the RGB space views of clear and haze images respectively, along with the corresponding decomposed views of the Y, U, and V components. It can be observed that the RGB haze images suffer from low contrast and local texture blur. The peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) between the clear and degraded images are 13.31dB and 71.92%, respectively. Converting the images to the YUV space and computing objective quality metrics for each component reveals that the pixel and structural differences caused by haze are predominantly manifested in the Y component. The PSNR and SSIM for the Y component are 14.67dB and 75.02%, respectively. Through visual comparison, it can be seen that the brightness of the Y channel in haze images is higher, presenting an overall misty appearance, while the UV color channels are less affected, evidenced by their higher objective metric values.

On the basis of the preceding analysis, the method³⁸ only remove the haze in Y channel and retain the UV channel of the original image. Such Y-channel-based strategy can accurately extract the haze distribution via the variations of luminance component, achieving superior dehazing performance. However, existing Y-channel

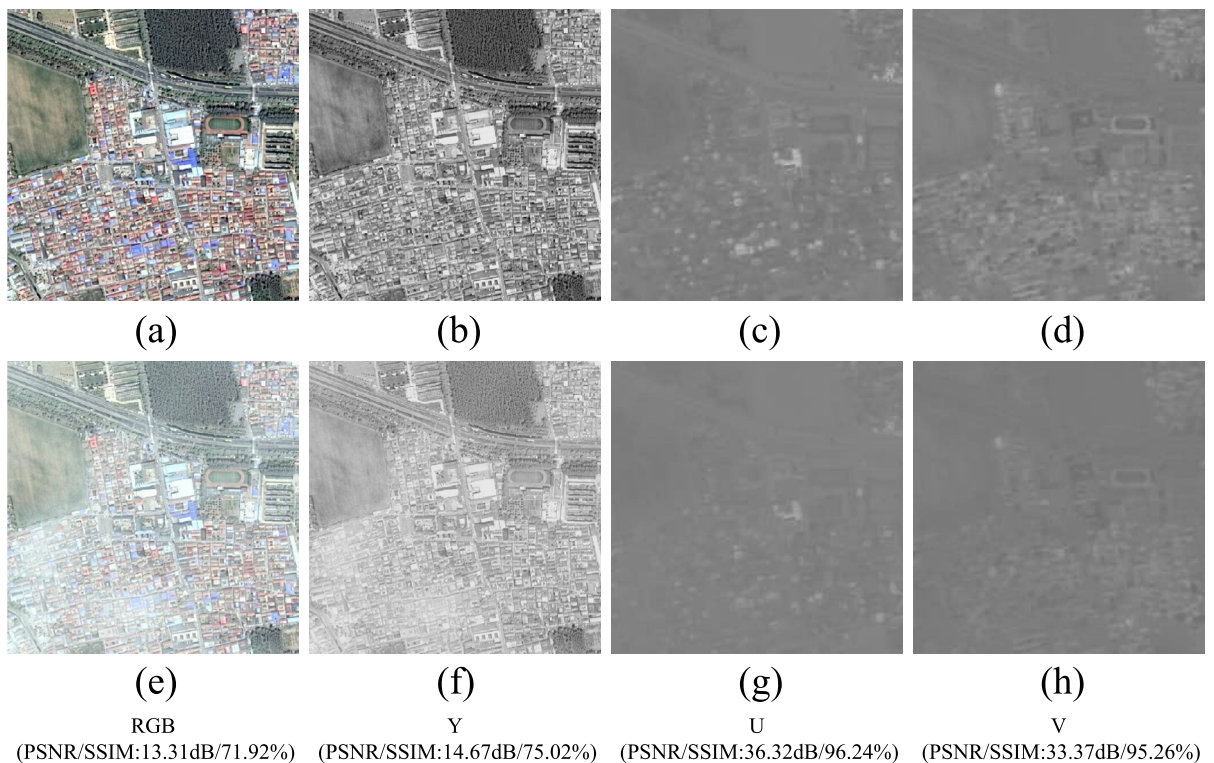


Fig. 2. Comparisons of clear image and RS haze image in RGB and Y, U, V spaces. (a) Clear image in RGB space; (b) Y component of clear image; (c) U component of clear image; (d) V component of clear image; (e) Haze image in RGB space; (f) Y component of haze image; (g) U component of haze image; (h) V component of haze image.

dehazing algorithms are not the optimal strategies. In detail, enhancing only Y channel and directly inheriting UV channel information from the original image may disrupt the synergy between Y, U, V channels. That will weaken the chromaticity feature and compress the image's contrast.

When we concatenate Y channel of the clear image presented in Fig. 3 (a) with the UV channel of the haze image presented in Fig. 3 (b) and convert it to an RGB image, we obtain a low-colorful haze-free image shown in Fig. 3 (c). It is evident that Fig. 3 (a) and Fig. 3 (c) have significant differences. To further visualise the distances, we perform color quantization on them and applied K-means clustering on the quantized results to extract the main 15 color points of the images. Compared to Fig. 3 (d), Fig. 3 (f) and Fig. 3 (h) have a more concentrated quantization distribution due to lower color richness and contrast. To further discuss the impacts of haze, we conduct a mathematical analysis of the atmospheric scattering model to elucidate the degradation characteristics observed in Y channel and UV channel. We consolidate Eqs.(4)-(6) into the matrix expression form:

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{7}$$

The expression of Eq. (1) can be reformulated as:

$$I_c = J_c t + A_c(1 - t), c \in \{R, G, B\} \tag{8}$$

As the scattering coefficient of the atmosphere β , exhibits minimal variation across channels, we can effectively disregard the distance component and consider t (i.e., $e^{-\beta d}$) as a constant. In the case of a color-balanced haze image, where the principle $A_R \approx A_G \approx A_B$ is consistently observed³⁹, upon substituting Eq. (7) into Eq. (8), we can derive the mapping relationship by rearranging the expressions:

$$\begin{aligned} I_Y &= (0.299J_R + 0.587J_G + 0.114J_B)t + A(1 - t) \\ &= J_Y t + A(1 - t), \\ I_U &= (-0.169J_R - 0.331J_G + 0.500J_B)t = J_U t, \\ I_V &= (0.500J_R - 0.419J_G - 0.081J_B)t = J_V t. \end{aligned} \tag{9}$$

Eq. (9) reveals that the atmospheric light A is exclusively associated with the luminance component Y , while the chrominance components U and V are subject to linear attenuation by the transmission map t and have no direct dependence on A . This result arises naturally from the mathematical derivation which is rooted in color space transformations and the atmospheric scattering model rather than being an a priori assumption. It aligns with the physical nature of haze: atmospheric light primarily contributes to the visual haze-like characteristics such as overall brightness obscuration and lacks significant chromaticity², and this is consistent with the observed behavior of chrominance components in hazy scenes as visualized in Figure 2. We then depict the conversion process between haze image I_{YUV} and its corresponding clear image J_{YUV} as follows:

$$I_{YUV} = [I_Y, I_U, I_V] = [J_Y t + A(1 - t), J_U t, J_V t] = J_{YUV} t + [A(1 - t), 0, 0] \tag{10}$$

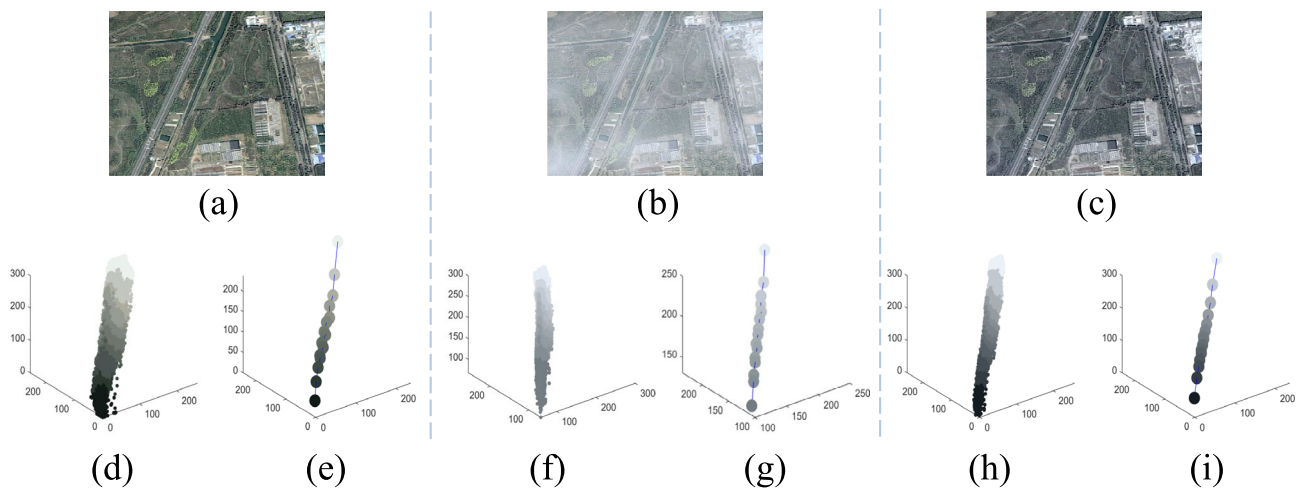


Fig. 3. Demonstration of color quantification differences. (a) Clear image; (b) Haze image; (c) Concatenated image; (d) Color quantization of clear image; (e) Color cluster of clear image; (f) Color quantization of haze image; (g) Color cluster of haze image; (h) Color quantization of concatenated image; (i) Color cluster of concatenated image.

As discussed, haze-induced alterations in luminance are global, driven by atmospheric light, whereas chrominance degradation stems from transmission-dependent attenuation. This analysis underscores that an effective YUV-based dehazing algorithm must address haze interference in the Y channel while accounting for the collaborative dynamics across the three channels.

RK method inspired dehazing network

As illustrated in Fig. 4, the proposed RKNet consists of two parts: haze removal and chrominance enhancement. The haze removal sub-net utilizes a pyramid structure, performing 2x upsampling and 2x downsampling on the original luminance component $I_Y \in \mathbb{R}^{H \times W \times 1}$, that aims to capture both local and global distribution features of the haze in Y domain. We embed the RK3 block in each scale branch to calculate the mapping relationship between state nodes in latent space, following which the pixel attention module²³ is employed to adaptively allocate weights for the output status of RK3 block. Finally, the feature information flow from the three branches is integrated using element-wise addition, resulting in the haze-free map $J'_Y \in \mathbb{R}^{H \times W \times 1}$.

For the attenuated chrominance component $I_{UV} \in \mathbb{R}^{H \times W \times 2}$, if we directly splice it with J'_Y , the channels feature may not be well fused. The converted RGB image will suffer from color degradation and visually tend towards greyscale. To address the issue, we introduce the chrominance enhancement (CE) block to balance the chromaticity component. Specifically, we embed the RK3 block and the channel attention module²³ into CE block to dynamically assign the weight ratios for the associated representations. The output of CE block is marked as $J'_{UV} \in \mathbb{R}^{H \times W \times 2}$, and we finally splice it with J'_Y and convert the spliced result back to RGB space. Overall, the proposed RKNet has a significantly concise structure without redundant branches. Such single-channel haze removal strategy greatly reduces the computational complexity of the algorithm and presents a promising direction for future lightweight dehazing research.

Converting RK method to component block

The explicit RK method is a common approach used in numerical ODE, the following formulas can describe arbitrary s stages of it mathematically:

$$y_{k+1} = y_k + \sum_{i=1}^s c_i G_i \tag{11}$$

$$\begin{cases} G_1 = hf(x_k, y_k) \\ G_i = hf\left(x_k + \lambda_i h, y_k + \sum_{j=1}^{i-1} \beta_{ij} G_j\right) \\ \sum_{i=1}^s c_i = 1, \quad \sum_{j=1}^{i-1} \beta_{ij} = \lambda_i \end{cases} \tag{12}$$

where G is the adaptively selected nonlinear representation unit; h is the stride length; c, β and λ are the constants, and can be calculated by the Taylor-expansion formula.

In the Runge-Kutta framework, higher-order methods generally reduce local truncation errors by incorporating more intermediate computations but at the expense of increased computational load^{40,41}. The 3rd-order Runge-Kutta (RK3) method emerges as an optimal compromise. It balances the need for accuracy, offering

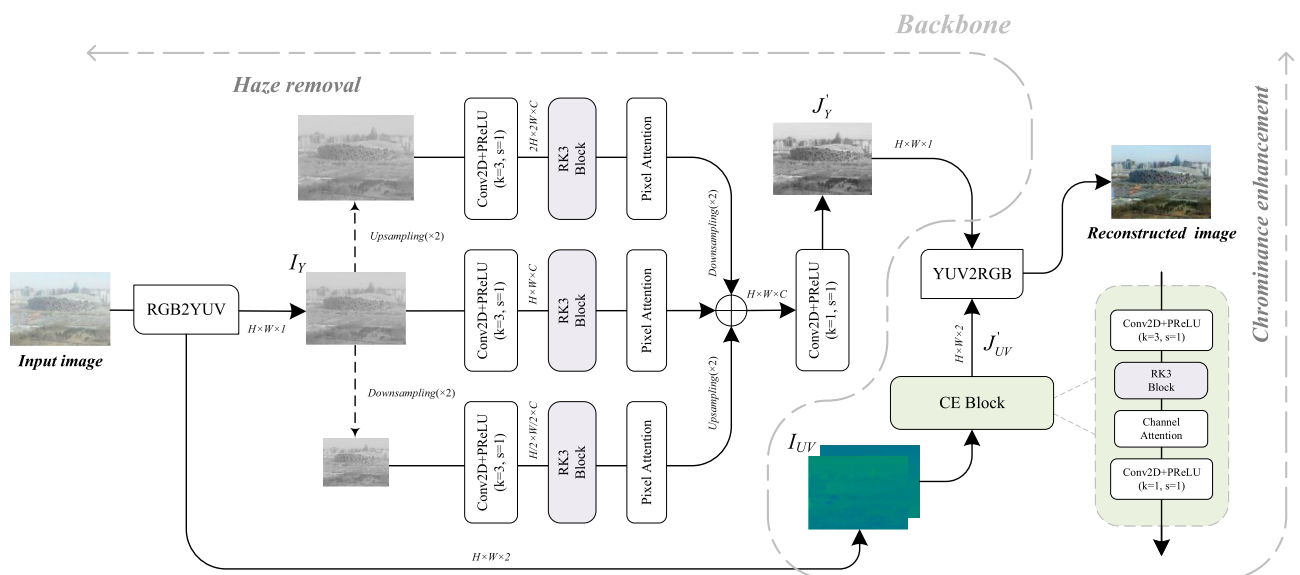


Fig. 4. The architecture of the proposed RKNet.

substantially lower truncation errors than lower-order alternatives, with computational efficiency, avoiding the excessive computational demands of higher-order schemes. This balance makes RK3 particularly well-suited for applications where both precision and computational feasibility are critical. Based on Eq. (11), the mathematical expression of the RK3 method can be formulated as:

$$y_{k+1} = y_k + \frac{1}{6}G_1 + \frac{2}{3}G_2 + \frac{1}{6}G_3 \tag{13}$$

$$\begin{cases} G_1 = hf(x_k, y_k) \\ G_2 = hf(x_k + \frac{h}{2}, y_k + \frac{1}{2}G_1) \\ G_3 = hf(x_k + h, y_k - G_1 + 2G_2) \end{cases} \tag{14}$$

We transform Eqs.(13, 14) into an intuitive network module. As depicted in Fig. 5, we denote the input state at the current moment as y_k , using y_{k+1} to represent the output state after a series of numerical computations. The specific structure of G will be further discussed in section "Discussion on the structure of G ".

Overall loss function

To optimize RKNet, we employ a combination of the \mathcal{L}_1 loss and the contrastive regularization loss⁹ as the objective function. \mathcal{L}_1 loss is commonly employed in low-level vision tasks, given N training sampling $\{I_i, J_i\}_{i=1}^N$, we describe it as:

$$\mathcal{L}_1 = \frac{1}{N} \sum_{i=1}^N \|J_i - Net(I_i)\| \tag{15}$$

Furthermore, we incorporate the contrastive regularization loss at the perceptual level to enhance the network's capability in extracting latent representations. Mathematically, it can be expressed as:

$$\mathcal{L}_{cr} = \sum_{i=1}^M \alpha_i F \{ \Phi_i(I), \Phi_i(J), \Phi_i[Net(I)] \} \tag{16}$$

$$F = \frac{D \{ \Phi(J), \Phi[Net(I)] \}}{D \{ \Phi(I), \Phi[Net(I)] \}} \tag{17}$$

here Φ denotes the fixed pre-trained VGG19 model adopted to extract the perceptual features; M is the number of the hidden layer; α_i is the trade-off weights; $D(x, y)$ indicates the \mathcal{L}_1 distance between x and y .

Based on the above considerations, the total loss function of RKNet is formulated as the following:

$$\mathcal{L}_T = \lambda_1 \mathcal{L}_1 + \lambda_2 \mathcal{L}_{cr} \tag{18}$$

here λ_1, λ_2 are the hyperparameters for balancing the pixel loss and contrastive regularization loss.

Experiments
Implementation details

In the training stage, we adopt 5000 pairs of synthetic haze images from RESIDE⁴² as the training data. All the inputs of the network are cropped to a size of $256 \times 256 \times 3$, and their pixel values are normalized. The training process is performed using PyTorch 1.9.0 on an NVIDIA GeForce RTX2070 GPU, and the network is trained for a total of 100 epochs. To optimize RKNet, we set λ_1 and λ_2 to 0.7 and 0.3, respectively. The trade-off weights α_i are determined using the approach proposed in⁹. We utilize the Adam optimizer with an initial learning

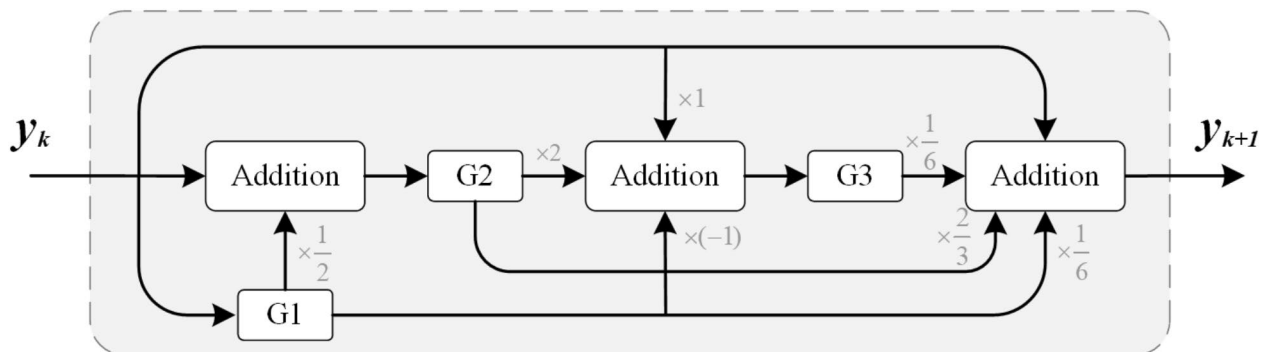


Fig. 5. The framework of RK3 block based on the 3rd-order explicit Runge-Kutta method. G1, G2 and G3 are the pre-defined G .

rate of 1×10^{-3} to update the network parameters throughout the training process. It's worth noting that the conversion from RGB to YUV spaces is fully reversible, ensuring the valuable feature of the input is preserved.

Comparison with SOTA algorithms

Evaluating on synthetic datasets

We compare the proposed RKNet with several SOTA haze removal algorithms involving CEEF⁴⁴, BCDP⁴⁵, FADE⁴⁶, PMD-Net⁴⁷, T-Net⁴⁸, and IHRNet⁴⁹. The experiments are performed on two synthetic hazy benchmarks, including RICE⁴³ and RESIDE⁴². The visual comparison on general daytime haze images is presented in Fig. 6. It is evident that CEEF⁴⁴ exhibits underexposure, BCDP⁴⁵ exposes color distortion issues in the sky area, and the results of FADE⁴⁶ appear visually dimmed. As the learning-based methods, PMD-Net⁴⁷, T-Net⁴⁸, and IHRNet⁴⁹ fail to effectively eliminate the haze effect.

Figure 6 provides the visual comparison on aerial haze images. It is evident that previous algorithms have primarily focused on general scenes, which may not be well-suited for aerial haze images. CEEF⁴⁴ is capable of eliminating the haze effect; however, excessive enhancement leads to a lack of visual hierarchy in the results. The contrast of BCDP's results⁴⁵ is suppressed, and the outputs of FADE⁴⁶ lose many color features. PMD-Net⁴⁷ and IHRNet⁴⁹ exhibit varying degrees of residual haze in a visually grey-masked manner. T-Net⁴⁸ performs poorly on aerial haze images, displaying evident color shift issues in its outputs. The visual comparison on general daytime haze images is presented in Fig. 7. It is evident that CEEF⁴⁴ exhibits underexposure, BCDP⁴⁵ exposes color distortion issues in the sky area, and the results of FADE⁴⁶ appear visually dimmed. As the learning-based methods, PMD-Net⁴⁷, T-Net⁴⁸, and IHRNet⁴⁹ fail to effectively eliminate the haze effect. Compared to these algorithms, RKNet can effectively eliminate haze interference while preserving the original brightness and color feature of the images. That is credited to the reasonable design of the improved dehazing strategy in YUV space.

To evaluate the performance of the algorithms, three evaluation metrics are utilized: PSNR, SSIM, and CIEDE2000. These metrics are responsible for measuring the pixel error, structure error, and color error of the images, respectively. Higher values of PSNR and SSIM, and lower values of CIEDE2000 indicate better algorithm performance. The metrics on RICE⁴³ as shown in Table 1. And the results of objective metrics on RESIDE⁴² are presented in Table 2. For comparison purposes, the best metrics are marked in bold and the second-best are underlined. Remote sensing images generally exhibit complex and diverse texture patterns, which makes the task of aerial image haze removal more challenging. Comparing Table 1 and Table 2, we notice that the overall performance of the algorithms on RICE is relatively lower. For T-Net⁴⁸, the metrics are not satisfactory due to its limited generalisation ability. In contrast, RKNet demonstrates better adaptability to various scenarios and achieves optimal performance in the quantitative evaluation.

Evaluating on real-world haze images

Despite being trained on synthetic datasets, RKNet demonstrates excellent generalization performance on real-world haze images. Figure 8 showcases the visual comparison of the algorithms. Although CEEF⁴⁴, BCDP⁴⁵, and FADE⁴⁶ can partially remove haze, their over-enhancement tends to distort color feature of the outputs. In the bottom row of Fig. 8, T-Net⁴⁸ and IHRNet⁴⁹ expose aberrant color blocks in the sky region, along with residual

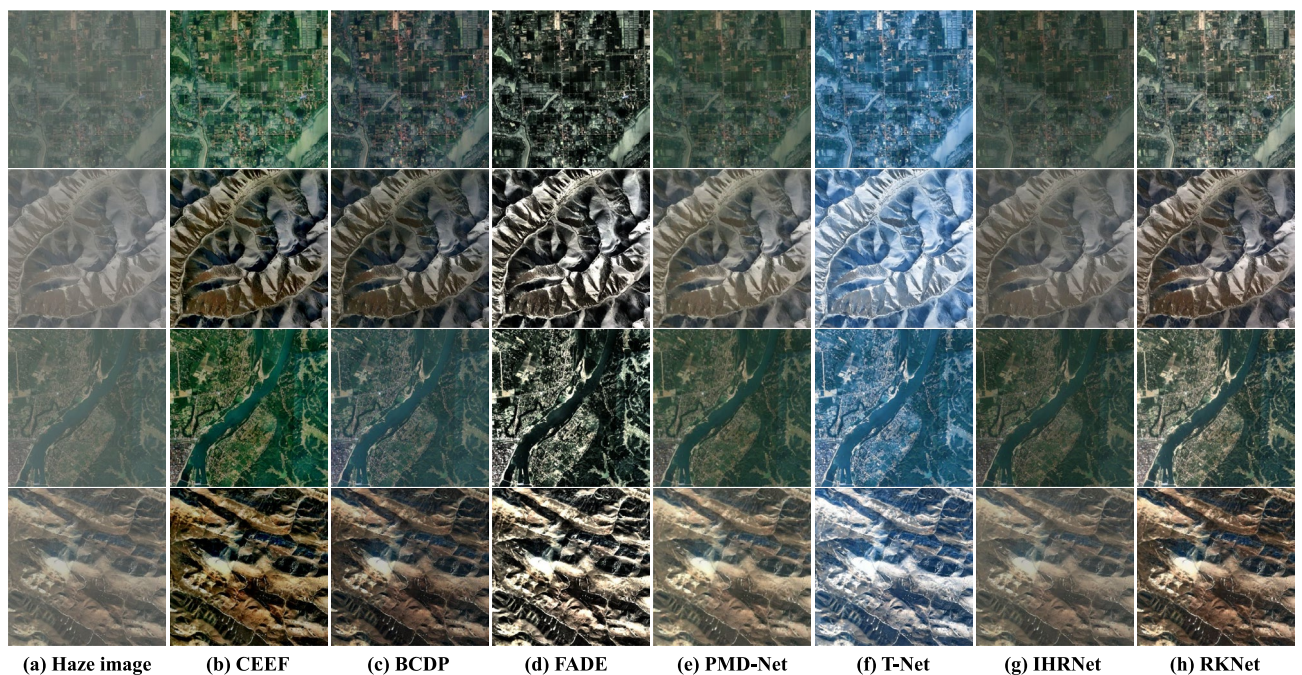


Fig. 6. The visual comparison on aerial haze images⁴³. CEEF⁴⁴, BCDP⁴⁵ and FADE⁴⁶ are the traditional dehazing methods; PMD-Net⁴⁷, T-Net⁴⁸ and IHRNet⁴⁹ are the learning-based dehazing methods.

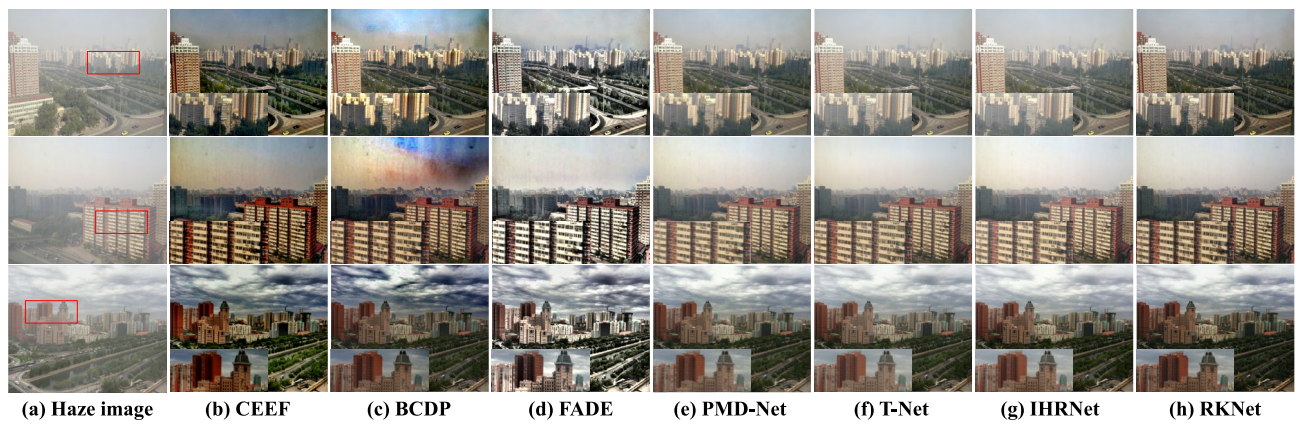


Fig. 7. The visual comparison on daytime haze images⁴². Please zoom in for a better illustration.

Methods	CEEF	BCDP	FADE	PMD-Net	T-Net	IHRNet	RKNet
PSNR(↑)	18.18	17.42	17.09	<u>20.82</u>	14.81	19.77	21.66
SSIM(↑)	0.8002	0.7806	0.7101	<u>0.8264</u>	0.5843	0.8135	0.8492
CIEDE2000(↓)	24.71	27.89	29.44	<u>20.21</u>	38.59	21.78	19.23

Table 1. The comparison of the metrics on aerial hazy dataset.

Methods	CEEF	BCDP	FADE	PMD-Net	T-Net	IHRNet	RKNet
PSNR(↑)	19.36	18.66	16.08	<u>32.35</u>	31.32	31.14	32.71
SSIM(↑)	0.7978	0.8483	0.7358	<u>0.9694</u>	0.9613	0.9684	0.9725
CIEDE2000(↓)	22.55	23.27	29.89	<u>13.83</u>	13.94	14.16	11.92

Table 2. The comparison of the metrics on daytime hazy dataset.

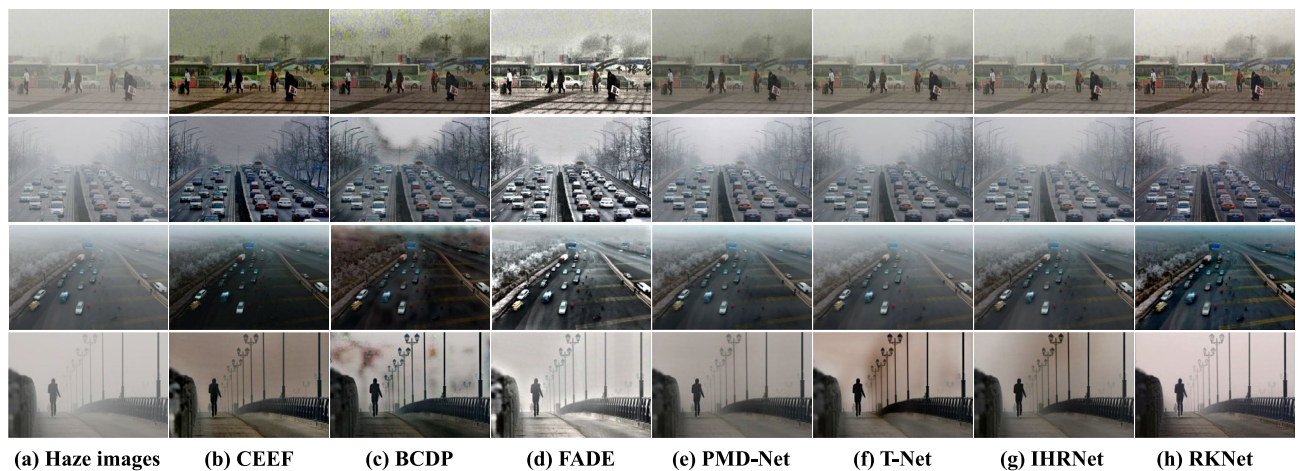


Fig. 8. The visual comparison on real-world haze images⁴².

Metrics	CEEF	BCDP	FADE	PMD-Net	T-Net	IHRNet	RKNet
MACs	30.92G	32.76G	<u>28.64G</u>	259.68G	37.67G	31.84G	20G
Parameters	1.77M	2.46M	<u>1.42M</u>	65.28M	2.55M	3.35M	0.07M

Table 3. The comparison of model complexity.

Image sizes	CEEF	BCDP	FADE	PMD-Net	T-Net	IHRNet	RKNet
960 × 540	0.1834	0.2427	<u>0.1727</u>	1.4606	0.7082	0.1921	0.1267
1280 × 720	0.3156	0.5003	<u>0.2989</u>	2.7038	1.2331	0.3308	0.1681
1920 × 1080	0.6454	0.8659	<u>0.5683</u>	4.9852	2.9673	0.6532	0.3137

Table 4. Comparison of time-consuming for different image sizes (in second).

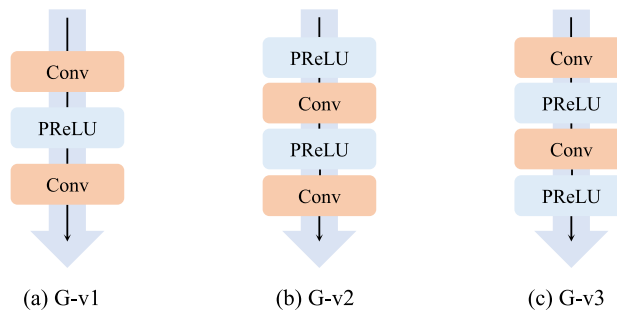


Fig. 9. The candidate structures of G , where 'Conv' denotes the convolutional layer and 'PReLU' represents the activation function.

Versions	G(-)	PSNR(\uparrow)	SSIM(\uparrow)	CIEDE2000(\downarrow)	MACs	Params
RKNet(v1)	G-v1	31.37	0.9593	12.99	20.00G	68.12k
RKNet(v2)	G-v2	32.28	0.9647	12.58	20.00G	68.12k
RKNet(v3)	G-v3	32.71	0.9725	11.92	20.00G	68.12k

Table 5. Metrics comparison of the candidate structures for G .

haze remaining. In comparison, RKNet not only comprehensively eliminates haze but also generates visually-natural results with fewer artifacts.

Time complexity analysis

In this section, we conduct experiments to analyze the time complexity of the methods. Table 3 presents the multiply-accumulate operations (MACs) and the number of parameters of the algorithms. Among the comparison models, RKNet has the lowest value of MACs and the fewest number of parameters. To further evaluate the computational efficiency, we then perform time-consuming experiment on a PC with an Intel(R) Core (TM) i5-9400 CPU@2.90GHz, 16GB RAM, and GPU@NVIDIA GeForce RTX2070. As is depicted in Table 4, we calculate the average time-consuming by the methods in handling 540p, 720p, and 1080p haze images. One can see that compared with the prior algorithms, RKNet demonstrates higher computational efficiency.

Discussion on the structure of G

As the adaptively selected nonlinear representation unit, there is a considerable space for exploring the structure of G . In this section, we aim to explore the performance of RKNet by experimenting with different designs of G .

Figure 9 illustrates the three structural schemes we considering for the ablation study. Each scheme consists of two convolutional layers, and the structure of the versions is defined by adjusting the position of the convolution and activation function. To determine the optimal strategy, we evaluate the candidates and calculate the objective indicators based on RESIDE⁴². As shown in Table 5, RKNet(v3) achieves the best metrics while maintaining the same computational complexity and number of parameters. Therefore, we incorporate G-v3 as the adaptively selected nonlinear representation unit in RKNet.

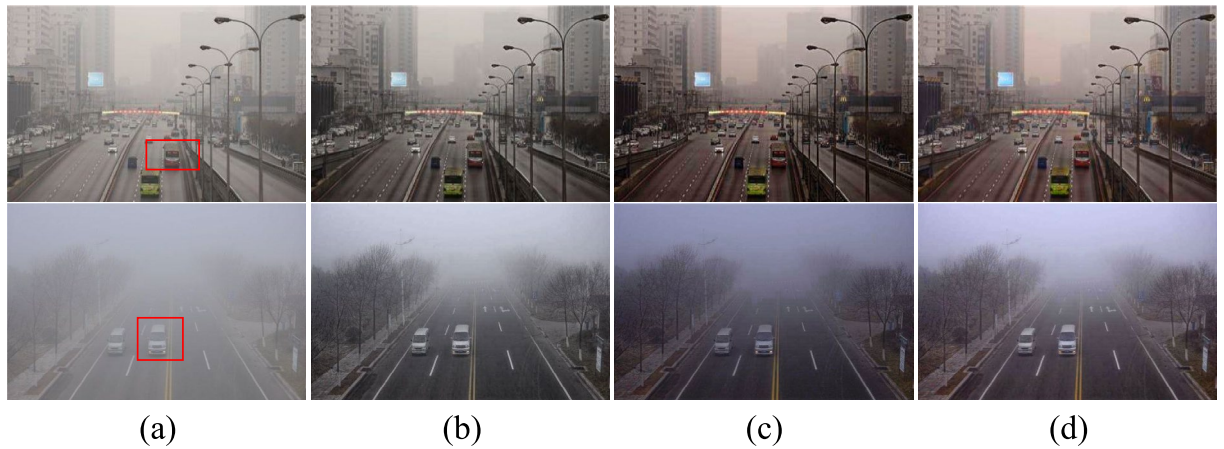


Fig. 10. The visual comparison of the dehazing strategies. (a) Haze images. (b) Only processing Y-channel. (c) Co-processing YUV channel. (d) Separate processing Y-channel and UV-channel.

Metrics	Scheme 1	Scheme 2	Scheme 3
BIQI(↓)	24.47	23.63	19.39
NIQE(↓)	3.78	4.26	3.02
SSEQ(↓)	26.71	26.98	24.16

Table 6. Metrics comparison of the candidate schemes.

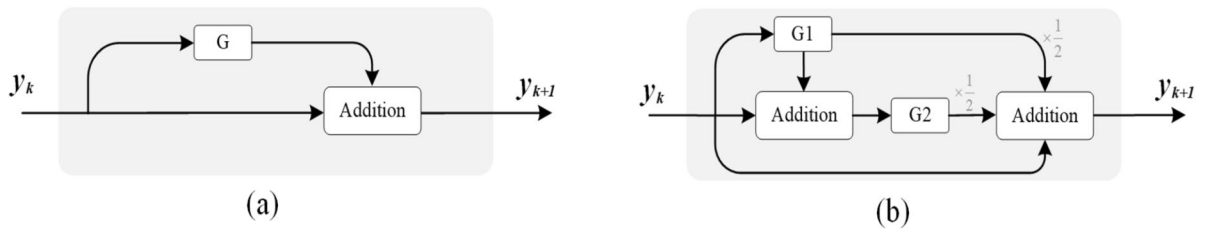


Fig. 11. The structure of the ODE-inspired components. (a) Residual block. (b) RK2 block.

Discussion on the dehazing strategy in YUV space

To explore the optimal dehazing strategy in YUV space, we conduct the ablation experiment in this section. Specifically, we mainly focus on these candidate schemes: 1) Enhancing Y-channel using the haze removal network and inheriting the UV-channel from the inputs. 2) Training the haze removal network to directly reconstruct haze-free images by taking Y, U, V channels as inputs. 3) Enhancing Y-channel using the haze removal network and refining the UV-channel via CE block, and then doing channel splicing.

Figure 10 presents the visual comparison of the dehazing strategies. In Fig. 10 (b), when we exclusively remove the haze in the Y-channel while directly inheriting the UV-channel from the original images, the visual results appear dim and colorless. Figure 10 (c) indicates that by cooperatively processing the Y, U, V channels as inputs for the haze removal network, the brightness features are compromised, leading to visually darker outputs. Figure 10 (d) demonstrates that the strategy of independently enhancing the Y and UV channels can generate visually more natural results with richer color information. We employed BIQI, NIQE, and SSEQ as objective metrics to quantitatively analyze the average performance of the three schemes on 200 real haze images, where lower metric values indicate better algorithmic performance. As shown in Table 6, the scheme that performs dehazing on the Y channel and collaborative enhancement on the UV channels achieved optimal performance in quantitative evaluations.

Comparative analysis of ODE-inspired component

To investigate the influence of different ODE-inspired architectures on network outputs, we conduct the ablation study by individually incorporating two more candidate blocks into the backbone model. As shown in Fig. 11 (a), the Residual block is a classic residual structure interpretable as a discrete approximation of ODE; the RK2

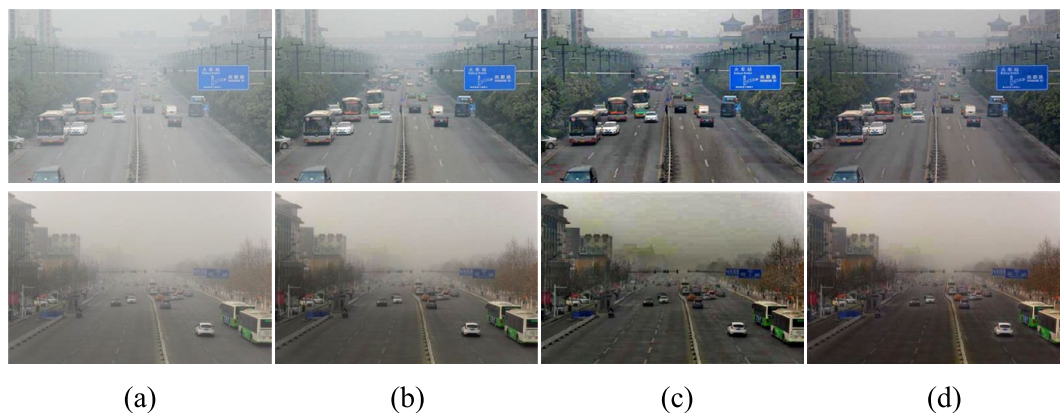


Fig. 12. Visual comparison of the components with different ODE-inspired structures in real-world haze images. (a) Haze images. (b) Results of the backbone model with embedded Residual blocks. (c) Results of the backbone model with embedded RK2 blocks. (d) The results of the backbone model with embedded RK3 blocks.



Fig. 13. Effect of RKNNet on the performance of YOLOv7⁵⁰. (a) Detection results of the haze images. (b) Detection results of the clear images reconstructed by RKNNet.

block's structure, presented in Fig. 11 (b), is constructed based on the second-order Runge-Kutta method to model continuous feature evolution via iterative updates.

Visual comparisons of these ODE-inspired components on real haze images are provided in Fig. 12. Specifically, Fig. 12 (b) shows that the network with embedded Residual blocks fails to effectively mitigate the effects of haze, resulting in visually obscure and blurry outputs. In Fig. 12 (c), although the network with embedded RK2 blocks reduces haze to some extent, the corresponding outputs exhibit noticeable artifacts, particularly in sky regions. Figure 12 (d) demonstrates that the proposed RK3 module not only effectively removes haze but also balances the synergistic relationship between brightness and chromaticity, yielding visually more natural reconstructed results.

Extension applications

Application in object detection

In hazy weather, the low-quality images captured by imaging facilities seriously affect the accuracy of high-level vision algorithms. Here, we present an experimental analysis to demonstrate the effectiveness of incorporating RKNNet in improving the accuracy of object detection algorithm under hazy conditions. To conduct this experiment, we utilized the popular object detection algorithm YOLOv7⁵⁰ as the baseline for our evaluation.

Figure 13 (a) shows the object detection results in hazy scenes, while Fig. 13 (b) displays the detection results after being processed by RKNNet. It is observed that the number of detected objects increased noticeably, indicating that the removal of haze boosted the visibility and made previously undetectable objects discernible. The results are particularly significant in high-density haze regions, where objects were previously completely obscured by the haze. Furthermore, the probabilities for dehazed objects exhibits overall improved, reflecting enhanced confidence of the algorithm in the detected object categories. This enhancement can be attributed to RKNNet effectively eliminating the visual artifacts caused by haze, resulting in more accurate object detection outcomes. The experimental results underscore the potential of RKNNet in addressing the challenges posed by hazy weather conditions in real-world object detection applications.



Fig. 14. The reconstructed examples of sandstorm images³⁹ using the fusion strategy, where (a) are the original sandstorm images, (b) are the color balanced images and (c) are the clear images enhanced by RKNNet.



Fig. 15. The enhanced examples of underwater images⁵² using RKNNet, where (a) are the underwater images and (b) are the clear images reconstructed by RKNNet.

Sandstorm image reconstruction

Fusion strategy⁵¹ has been proven to be feasible for sandstorm image enhancement task. Following that line, we conduct a simple color equalization pre-processing⁵¹ and explore the feasibility of utilizing RKNNet as the dust removal module in fusion strategy-based sand-dust image enhancement method. Given the sandstorm images as shown in Fig. 14 (a), we can see from Fig. 14 (b) that the color equalization method effectively balances the hue of the images and eliminates the interference of sand, but the dust still exists. In Fig. 14 (c), RKNNet can further improve the clarity of the images by eliminating dust interference while retaining the local details.

Underwater image enhancement

Although the proposed algorithm is designed for haze image reconstruction, it has potential applications in complex underwater environments. As shown in Fig. 15, we provide several enhanced examples of underwater images. The top row is the original underwater images and the bottom row is the enhanced results using the proposed method. It is evident that RKNNet successfully eliminates the haziness in the images and restores the color characteristics of the scene. From the aesthetic perspective, RKNNet can effectively improve the visual quality of the degraded underwater images.

Conclusion

In this paper, we first analysed the weaknesses in current YUV-based dehazing methods and introduced an improved haze removal strategy in YUV space. Following the optimised strategy, we proposed a novel image dehazing network called RKNNet, which consists of Y domain haze removal and UV domain chrominance enhancement two parts. Based on the 3rd-order Runge-Kutta method, we designed a RK3 block and integrated it into RKNNet to reduce parameter redundancy and computational costs. The experimental results on synthetic and real-world benchmarks demonstrate that RKNNet outperforms the SOTA algorithms both qualitatively and quantitatively. Moreover, the proposed algorithm has the potential to improve the performance of object

detection model and also generalizes well in other low-level image processing tasks, such as sandstorm image reconstruction and underwater image enhancement.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 21 June 2025; Accepted: 26 September 2025

Published online: 03 November 2025

References

- Singh, D. & Kumar, V. Comprehensive survey on haze removal techniques. *Multimed. Tools Appl.* **77**, 9595–9620 (2018).
- He, K., Sun, J. & Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**, 2341–2353 (2010).
- Bui, T. M. & Kim, W. Single image dehazing using color ellipsoid prior. *IEEE Trans. Image Process.* **27**, 999–1009 (2017).
- Kaur, M., Singh, D., Kumar, V. & Sun, K. Color image dehazing using gradient channel prior and guided l0 filter. *Inf. Sci.* **521**, 326–342 (2020).
- Wan, D. et al. Yolo-hr: Improved yolov5 for object detection in high-resolution optical remote sensing images. *Remote Sens.* **15**, 614 (2023).
- Xiao, J. et al. Tiny object detection with context enhancement and feature purification. *Expert Syst. Appl.* **211**, 118665 (2023).
- Dong, X. et al. Attention-based multi-level feature fusion for object detection in remote sensing images. *Remote Sens.* **14**, 3735 (2022).
- 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.* **25**, 5187–5198 (2016).
- Wu, H. et al. Contrastive learning for compact single image dehazing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10551–10560 (2021).
- Si, Y., Yang, F. & Chong, N. A novel method for single nighttime image haze removal based on gray space. *Multimed. Tools Appl.* **81**, 43467–43484 (2022).
- Bai, H., Pan, J., Xiang, X. & Tang, J. Self-guided image dehazing using progressive feature fusion. *IEEE Trans. Image Process.* **31**, 1217–1229 (2022).
- Fang, F., Wang, T., Wang, Y., Zeng, T. & Zhang, G. Variational single image dehazing for enhanced visualization. *IEEE Trans. Multimed.* **22**, 2537–2550 (2020).
- Zhang, L. & Wang, S. Dense haze removal based on dynamic collaborative inference learning for remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **60**, 1–16 (2022).
- Yadav, A. C., Bose, S. & Kolekar, M. H. Image enhancement by defogging utilizing scene specific dark channel prior. In *2024 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, 1–6 (IEEE, 2024).
- Liu, J., Liu, W., Sun, J. & Zeng, T. Rank-one prior: Toward real-time scene recovery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14802–14810 (2021).
- Liu, J., Liu, R. W., Sun, J. & Zeng, T. Rank-one prior: Real-time scene recovery. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**, 8845–8860 (2023).
- Ling, Z., Fan, G., Wang, Y. & Lu, X. Learning deep transmission network for single image dehazing. In *2016 IEEE international conference on image processing (ICIP)*, 2296–2300 (IEEE, 2016).
- Kim, S. E., Park, T. H. & Eom, I. K. Fast single image dehazing using saturation based transmission map estimation. *IEEE Trans. Image Process.* **29**, 1985–1998 (2019).
- Yang, D. & Sun, J. Proximal dehaze-net: A prior learning-based deep network for single image dehazing. In *Proceedings of the european conference on computer vision (ECCV)*, 702–717 (2018).
- Ren, W. et al. Single image dehazing via multi-scale convolutional neural networks. In *14th European Conference on Computer Vision, ECCV 2016*, 154–169 (Springer Verlag, 2016).
- Su, Y. Z., He, C., Cui, Z. G., Li, A. H. & Wang, N. Physical model and image translation fused network for single-image dehazing. *Pattern Recognit.* **142**, 109700 (2023).
- Zhang, H. & Patel, V. M. Densely connected pyramid dehazing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3194–3203 (2018).
- 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*, 11908–11915 (2020).
- Wang, Y. et al. Ucl-dehaze: Toward real-world image dehazing via unsupervised contrastive learning. *IEEE Trans. Image Process.* **33**, 1361–1374 (2024).
- Su, Y. et al. Real scene single image dehazing network with multi-prior guidance and domain transfer. *IEEE Transactions on Multimedia* 1–16 (2025).
- Zhang, X., Wang, T., Wang, J., Tang, G. & Zhao, L. Pyramid channel-based feature attention network for image dehazing. *Comput. Vis. Image Underst.* **197**, 103003 (2020).
- Wei, J. et al. Shallow feature matters for weakly supervised object localization. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 5993–6001 (2021).
- Weinan, E. A proposal on machine learning via dynamical systems. *Commun. Math. Stat.* **1**, 1–11 (2017).
- He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778 (2016).
- Chang, B. et al. Reversible architectures for arbitrarily deep residual neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, 2811–2818 (2018).
- Lu, Y., Zhong, A., Li, Q. & Dong, B. Beyond finite layer neural networks: Bridging deep architectures and numerical differential equations. In *International Conference on Machine Learning*, 3276–3285 (PMLR, 2018).
- Xie, S., Girshick, R., Dollár, P., Tu, Z. & He, K. Aggregated residual transformations for deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1492–1500 (2017).
- Bai, Y., Liu, M., Yao, C., Lin, C. & Zhao, Y. Ode-inspired image denoiser: An end-to-end dynamical denoising network. In *Pattern Recognition and Computer Vision: 4th Chinese Conference, PRCV 2021, Beijing, China, October 29–November 1, 2021, Proceedings, Part III* 4, 261–274 (Springer, 2021).
- Yang, Y. et al. Dynamical system inspired adaptive time stepping controller for residual network families. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 6648–6655 (2020).
- He, X. et al. Ode-inspired network design for single image super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1732–1741 (2019).

36. Shen, J., Li, Z., Yu, L., Xia, G.-S. & Yang, W. Implicit euler ode networks for single-image dehazing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 218–219 (2020).
37. BT, R. I.-R. et al. Studio encoding parameters of digital television for standard 4: 3 and wide-screen 16: 9 aspect ratios. *International Radio Consultative Committee International Telecommunication Union, Switzerland, CCIR Rep* (2011).
38. Wang, W., Wang, A. & Liu, C. Variational single nighttime image haze removal with a gray haze-line prior. *IEEE Trans. Image Process.* **31**, 1349–1363 (2022).
39. Si, Y., Yang, F., Guo, Y., Zhang, W. & Yang, Y. A comprehensive benchmark analysis for sand dust image reconstruction. *J. Vis. Commun. Image Represent.* **89**, 103638 (2022).
40. Zhu, M., Chang, B. & Fu, C. Convolutional neural networks combined with runge-kutta methods. *Neural Comput. Appl.* **35**, 1629–1643 (2023).
41. Casas-Ordaz, A., Oliva, D., Navarro, M. A., Ramos-Michel, A. & Pérez-Cisneros, M. An improved opposition-based runge kutta optimizer for multilevel image thresholding. *J. Supercomput.* **79**, 17247–17354 (2023).
42. Li, B. et al. Benchmarking single-image dehazing and beyond. *IEEE Trans. Image Process.* **28**, 492–505 (2018).
43. Zhou, X., Lin, D. & Junyi, L. Rice: A dataset and baseline for cloud removal in remote sensing images. *J. Comb. Math. Comb. Comput.* **120**, 107–124 (2024).
44. Liu, X., Li, H. & Zhu, C. Joint contrast enhancement and exposure fusion for real-world image dehazing. *IEEE Trans. Multimed.* **24**, 3934–3946 (2022).
45. Zhao, X. Single image dehazing using bounded channel difference prior. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 727–735 (2021).
46. Li, Z. et al. Fast region-adaptive defogging and enhancement for outdoor images containing sky. In *2020 25th International Conference on Pattern Recognition (ICPR)*, 8267–8274 (IEEE, 2021).
47. Ye, T. et al. Perceiving and modeling density for image dehazing. In *Computer Vision—ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XIX*, 130–145 (Springer, 2022).
48. Zheng, L., Li, Y., Zhang, K. & Luo, W. T-net: Deep stacked scale-iteration network for image dehazing. *IEEE Trans. Multimed.* **25**, 6794–6807 (2023).
49. Si, Y., Li, C. & Yang, F. Ihr-net: image haze removal network based on improved yuv spatial strategy. *Meas. Sci. Technol.* **36**, 0161b7 (2025).
50. Wang, C.-Y., Bochkovskiy, A. & Liao, H.-Y. M. Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7464–7475 (2023).
51. Si, Y., Yang, F. & Liu, Z. Sand dust image visibility enhancement algorithm via fusion strategy. *Sci. Rep.* **12**, 13226 (2022).
52. Li, C. et al. An underwater image enhancement benchmark dataset and beyond. *IEEE Trans. Image Process.* **29**, 4376–4389 (2020).

Author contributions

Y.S. conducted the experiments and wrote the draft; J.C. prepared the Figures and Tables; W.R. carried out checks and revisions on the manuscript; C.L. analyzed the feasibility of the algorithm and managed the project. All authors reviewed the manuscript.

Funding

This work is supported by the Wuxi University Research Start-up Fund for High-level Talents(2025r028).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to C.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025